

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:13:43 ; Search time 38 Seconds

(without alignments)

1511.343 Million cell updates/sec

Title: US-10-070-241H-1

perfect score:

Sequence: 1 MQALNITPEQFSRIILRDHNL PHIALFRSFLAFNSPLDSGH 431

Scoring table: BLOSUM62

Gapop 10 0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```

Database :
A_Geneseq_101002:*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query |     | Length | DB       | ID | Description        |
|------------|-------|-------|-----|--------|----------|----|--------------------|
|            |       | Match | %   |        |          |    |                    |
| 1          | 2235  | 100.0 | 431 | 21     | AAV94993 |    | Human secreted pro |
| 2          | 2235  | 100.0 | 431 | 21     | AA564289 |    | Human GTP-binding  |
| 3          | 2235  | 100.0 | 431 | 22     | AAH74773 |    | Human G protein-co |
| 4          | 2235  | 100.0 | 431 | 22     | AA848963 |    | Human G protein-co |
| 5          | 2235  | 100.0 | 431 | 23     | AAU97735 |    | Human TGR346 poly  |
| 6          | 2235  | 100.0 | 431 | 23     | AAU76035 |    | Human neuropeptide |
| 7          | 2232  | 99.9  | 431 | 21     | AAV71309 |    | Human orphan G pro |
| 8          | 2232  | 99.9  | 431 | 21     | AA802843 |    | Human G protein-co |
| 9          | 2226  | 99.6  | 431 | 21     | AAU02953 |    | Human G protein-co |
| 10         | 1914  | 85.6  | 433 | 23     | AAU97740 |    | Mousn TGR346a poly |

## ALIGNMENT'S

## RESULT 1

AA94993

ID AAY94993 standard; Protein; 431 AA.

AC AAY94993;

DT 19-JUN-2000 (first entry)

Human secreted protein vc38\_1, SE0 ID NO:26.

Human; secreted protein; cancer; tumour; cardiovascular disorder;  
XX  
blood disorders; haemophilia; autoimmune disease; diabetes; inflan-  
KW  
infection; fungal; bacterial; viral; HIV; allergy; arthritis;  
KW  
neurodegenerative disease; asthma; contraceptive.  
KW

OS Homo sapiens.

AA WO200011015-A1

02-MAR-2000.

XX  
PF: 24-AUG-1999: 99WO-US19351.

XX  
EP 24 AUG-1998 0805-0097638

PK 24-AUG-1998; 98115-01997659  
PK 014 SEP-1998; 98115-01199618

28-SEP-1998; 98US-0102092.  
25-NOV-1998; 98US-0109978

|    |              |               |
|----|--------------|---------------|
| FR | 23-DEC-1998; | 98US-0113645. |
| DP | 23-DEC-1998. | 097C-0113645. |

PR 23 AUG-1999, 99US-0379246.

PA (ALPH-) ALPHAGENE INC.

\_\_\_\_\_

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;  
 XX WPI; 2000-224657/19.  
 XX New secreted or transmembrane proteins and polynucleotides encoding  
 PT them, useful for treating neurodegenerative disorders, autoimmune  
 PT diseases and cancer.  
 XX Claim 35; Page 284-285; 357pp; English.  
 XX The invention relates to 40 human secreted proteins (AA94981-Y95020),  
 CC and cDNA sequences encoding them (AA23423-A23462). The secreted  
 CC proteins of the invention include those that are thought to be only  
 CC partially secreted, i.e., transmembrane proteins. The proteins of the  
 CC invention may exhibit one or more activities selected from the following:  
 CC cytokine activity; cell proliferation; differentiation; immune  
 CC modulation; haematopoiesis regulation; tissue growth activity;  
 CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic  
 CC and thrombolytic activity; anti-inflammatory activity; and tumour  
 CC inhibition activity. The proteins may be administered to patients as  
 CC vaccines, and the nucleotides may be used as part of a gene therapy  
 CC regime, diseases or conditions that may be treated using the proteins or  
 CC nucleotides of the invention include autoimmune diseases; genetic  
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,  
 CC fungal and viral infections; especially HIV; multiple sclerosis;  
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;  
 CC insulin dependent diabetes mellitus; and allergic reactions such as  
 CC asthma and anaemia. They may also be used for treating wounds, burns,  
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic  
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may  
 CC additionally be useful as contraceptives. Nucleic acid sequences of the  
 CC invention may be used in chromosome mapping, and as a source of  
 CC diagnostic primers and probes. The present sequence represents one of the  
 CC 40 proteins of the invention.

XX Sequence 431 AA:

Query Match 100.0%; Score 2235; DB 21; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-237;  
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLAIVLTVGLIFALAL 60  
 DB 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLAIVLTVGLIFALAL 60

QY 61 FGNALFYVYVTRSKAMRTVTNIFICSLASDLLITFEFCIPVTMLQNSDNLGGAFCCKM 120  
 DB 61 FGNALFYVYVTRSKAMRTVTNIFICSLASDLLITFEFCIPVTMLQNSDNLGGAFCCKM 120

QY 121 VPEVQSTAVVTEILLMTCTIAVERHOGVLVHPFKMKWQYTNRRAFVWLVAVIVGSPM 180  
 DB 121 VPEVQSTAVVTEILLMTCTIAVERHOGVLVHPFKMKWQYTNRRAFVWLVAVIVGSPM 180

QY 181 WHVQOQLEIKYDFLYEKEHICCLEEWTSPVHQKIYTFILVILFLPLVWMLILYSKIGYE 240  
 DB 181 WHVQOQLEIKYDFLYEKEHICCLEEWTSPVHQKIYTFILVILFLPLVWMLILYSKIGYE 240

QY 241 LWIKKRVGDSVLRTIHGKEMSKIAKKKRAVIMVYVVALFVAVWAPFHVHMMIEYSN 300  
 DB 241 LWIKKRVGDSVLRTIHGKEMSKIAKKKRAVIMVYVVALFVAVWAPFHVHMMIEYSN 300

QY 301 FEKEYDDVTIKMIFALVQIHGFSNSICNPILVAFNENPKKVLISAVCYCIVNKTFSQAQ 360  
 DB 301 FEKEYDDVTIKMIFALVQIHGFSNSICNPILVAFNENPKKVLISAVCYCIVNKTFSQAQ 360

QY 361 RHGNSGITMWRKAKFSLPFPNVEETKTEAFSDVNIYVKLCEQTEERKKLKRHLALPSE 420  
 DB 361 RHGNSGITMWRKAKFSLPFPNVEETKTEAFSDVNIYVKLCEQTEERKKLKRHLALPSE 420

QY 421 LAENSPDLSGHH 431  
 DB 421 LAENSPDLSGHH 431

DB 421 LAENSPDLSGHH 431

RESULT 2  
 AAG64289  
 ID AAG64289 standard; Protein: 431 AA.  
 XX AC AAG64289;  
 XX DT 21-SEP-2001 (first entry)  
 XX DE Human GTP-binding protein-coupled receptor GPRV11.  
 XX KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;  
 XX KW muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;  
 XX KW G-protein.  
 XX OS Homo sapiens.  
 XX PN W0200148189-A1.  
 XX DB 05-JUL-2001.  
 XX DT 28-DEC-2000; 2000W0-JP09409.  
 XX PR 28-DEC-1999; 94JP-0375152.  
 XX PR 31-MAR-2000; 2000JP-0101339.  
 XX PR 23 MAY 2000; 2000JP-0155978.  
 XX PA (HELI-) HELIX RES INST.  
 XX PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Sawa M;  
 XX PI Sugiyama T;  
 XX DR WPI; 2001-425663/45.  
 XX DR N-PSDB; AAH49502.  
 XX PT Family of guanosine triphosphate binding protein coupled receptors and  
 XX PT genes encoding them for treatment and prevention of diseases associated  
 XX PT with these receptors.  
 XX PS Claim 1; Pages 59-62; 137pp; Japanese.

The present sequence is the protein sequence for a human guanosine  
 triphosphate (GTP)-binding protein-coupled receptor. The receptor is  
 useful for the investigation, diagnosis, treatment and prevention of  
 diseases associated with GTP-binding protein-coupled receptors, including  
 neurological, circulatory, digestive system, immune system, muscle and  
 urinary system disorders. GTP-binding proteins are also known as  
 G-proteins.

Query Match 100.0%; Score 2235; DB 22; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-237;  
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLAIVLTVGLIFALAL 60  
 DB 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLAIVLTVGLIFALAL 60

QY 61 FGNALFYVYVTRSKAMRTVTNIFICSLASDLLITFEFCIPVTMLQNSDNLGGAFCCKM 120  
 DB 61 FGNALFYVYVTRSKAMRTVTNIFICSLASDLLITFEFCIPVTMLQNSDNLGGAFCCKM 120

QY 121 VPEVQSTAVVTEILLMTCTIAVERHOGVLVHPFKMKWQYTNRRAFVWLVAVIVGSPM 180  
 DB 121 VPEVQSTAVVTEILLMTCTIAVERHOGVLVHPFKMKWQYTNRRAFVWLVAVIVGSPM 180

QY 181 WHVQOQLEIKYDFLYEKEHICCLEEWTSPVHQKIYTFILVILFLPLVWMLILYSKIGYE 240  
 DB 181 WHVQOQLEIKYDFLYEKEHICCLEEWTSPVHQKIYTFILVILFLPLVWMLILYSKIGYE 240

DB 181 WHVQOQLEIKYDFLYEKEHICCLEEWTSPVHQKIYTFILVILFLPLVWMLILYSKIGYE 240

[illegible]

PT infections, cancer, neurological disorders and asthma

PS Claim 1: Page 29-40; 31pp; English.

XX The invention relates to the human G protein-coupled receptor AXOR16  
 CC (AAH48964) to DNA encoding AXOR16 (AAH87690), and to AXOR16 fragments  
 CC and variants. Like all G protein-coupled receptors, AXOR16 has 7  
 CC putative transmembrane domains and is involved in signal transduction.  
 CC AXOR16 has homology with GAD65 morhua neurotrophin (NPYRB) F and  
 CC homology/ structural similarity with human neurotrophin Y (NPYRYA).  
 CC The human AXOR16 gene is located on chromosome 11q22.2. The invention  
 CC also relates to expression vectors and host cells comprising AXOR16 DNA,  
 CC to recombinant expression of AXOR16, and to an AXOR16-specific antibody.  
 CC AXOR16 proteins and nucleotides may be used to treat a wide variety of  
 CC disorders including bacterial, fungal, protozoal and viral infections,  
 CC particularly HIV-1 or HIV-2 infections; pain; cancers; benign prostatic  
 CC hypertrophy; diabetes; obesity; anorexia; bulimia; osteoporosis; asthma;  
 CC allergies; urinary retention; acute heart failure; hypotension;  
 CC hypertension; angina pectoris; myocardial infarction; stroke; glaucoma;  
 CC migraine; vomiting; psychotic and neurological disorders such as  
 CC anxiety, schizophrenia, manic depression, depression, delirium, dementia,  
 CC and severe mental retardation, and dyskinesias, such as Parkinson's  
 CC disease, Huntington's disease or Gilles de la Tourette's syndrome.  
 CC AXOR16 proteins, and nucleotides are useful as vaccines, and AXOR16  
 CC proteins, nucleotides and antibodies may be used in screening compounds  
 CC for their ability to modulate AXOR16 activity or expression. AXOR16  
 CC proteins are also useful for inducing an immunological response in a  
 CC mammal against the above diseases, for antibody production, and for  
 CC identifying membrane bound or soluble receptors for AXOR16. AXOR16  
 CC nucleotides are also useful as diagnostic reagents and in chromosome  
 CC localisation and tissue expression studies. The present sequence  
 CC represents human AXOR16.

XX Sequence 431 AA:

Query Match 100.0%; Score 2345; DB 22; Length 431.  
 Best Local Similarity 100.0%; Pred. No. 2.8e-237;  
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQFSKLLRDHNLTRFOPTALYRLPLVYTPPELPGRAKLAIVLTGVLFALAL 60  
 DB 1 MQALNITPEQFSKLLRDHNLTRFOPTALYRLPLVYTPPELPGRAKLAIVLTGVLFALAL 60

QY 61 FGNALVYVYVTRSKAMRTVTNIFICSLALSLLITFFCIPVTMLQNSIDNWLGAFICCKM 120  
 DB 61 FGNALVYVYVTRSKAMRTVTNIFICSLALSLLITFFCIPVTMLQNSIDNWLGAFICCKM 120

QY 121 VPPVQSTAVVTEILLTWCIAVERHQGLVHPFKMKWQYTNRRAPFTMLGVVWLVAVIVGSPM 180  
 DB 121 VPPVQSTAVVTEILLTWCIAVERHQGLVHPFKMKWQYTNRRAPFTMLGVVWLVAVIVGSPM 180

QY 181 WHVQQLLEKYDFLEYKEHICCLFEWTSPIHQKIYTFILVILFLLPLVMVLLYSKIGYE 240  
 DB 181 WHVQQLLEKYDFLEYKEHICCLFEWTSPIHQKIYTFILVILFLLPLVMVLLYSKIGYE 240

QY 241 LWIKKRVGDSVLTIRHKEMSKIAKKKRAVIMMVTVALFVGCWAPFHVHMMIRYSN 300  
 DB 241 LWIKKRVGDSVLTIRHKEMSKIAKKKRAVIMMVTVALFVGCWAPFHVHMMIRYSN 300

QY 301 FEKEYDDVTIKMIFAIQVIGFSNSICNPVYAFNENPKKVLISAVCYCIVNKTFSQAQ 360  
 DB 301 FEKEYDDVTIKMIFAIQVIGFSNSICNPVYAFNENPKKVLISAVCYCIVNKTFSQAQ 360

QY 361 RHGNSGTTMWRKAKFSLRNPVEFTKGFAPFSNGNIEVKTCFQTEEEKKKLKHLLALPSE 420  
 DB 361 RHGNSGTTMWRKAKFSLRNPVEFTKGFAPFSNGNIEVKLCFQTEEEKKKLKHLLALPSE 420

QY 421 LAFNSPLDSCH 431  
 DB 421 LAFNSPLDSCH 431

RESULT 5

AAU97735

XX ID AAU97735 standard; Protein; 431 AA.

XX AC AAU97735;

XX DT 27-AUG-2002 (first entry)

XX DE Human TGR346 polypeptide.

XX KW Human; TGR346; receptor; G-protein coupled receptor; GPCR; TGR;

XX KW TGR-associated disorder; signal transduction; renal failure; nephritis;

XX KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;

XX KW diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;

XX KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;

XX KW circadian rhythm.

XX OS Homo sapiens.

XX XX WC000242458-A2.

XX PD 30-MAY-2002.

XX PP 21-NOV-2001; 2001WO-0543404.

XX PP 22-NOV-2000; 2000US-252841P.

XX PP 22-DEC-2000; 2000US-257636P.

XX PP 12-JAN-2001; 2001US-261377P.

XX PP 28-MAR-2001; 2001US-279554P.

XX PP 29-MAR-2001; 2001US-280696P.

XX XX (TULIA-) TULAHIAK INC.

XX XX Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupta JS;

XX XX WPI: 2002-453633/49.

XX XX N-PSDB; ABK86287.

XX PT New isolated G-protein coupled receptor polypeptide, termed TGR, for  
 PT diagnosis and treatment of diseases such as renal failure, nephritis,  
 PT hypothyroidism, diabetes insipidus, and disturbances of thirst and  
 PT sleep.

XX PS Claim 33; Page 74; 98pp; English.

XX CC The invention relates to a G-protein coupled receptor polypeptide (GPCR),  
 CC termed TGR, and its associated nucleic acid. The sequences of the  
 CC invention are useful for identifying a compound that modulates signal  
 CC transduction and for identifying a mammal having a TGR-associated  
 CC disorder. The proteins and nucleic acids are useful in diagnosis and  
 CC treatment of diseases or conditions such as renal failure, nephritis,  
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,  
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,  
 CC sleep, temperature regulation, appetite, blood pressure or any other  
 CC syndrome or disease associated with the hypothalamus. The sequences can  
 CC be used in regulation of circadian rhythms, for use as genetic markers  
 CC for the identification of mutations associated with diseases resulting  
 CC from GPCR inactivation in particular cell types and for identification of  
 CC modulators of GPCR signal transduction. This sequence represents the  
 CC human TGR346 polypeptide.

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 2335; DB 23; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.8e-237;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQFSKLLRDHNLTRFOPTALYRLPLVYTPPELPGRAKLAIVLTGVLFALAL 60

DB 1 MQALNITPEQFSKLLRDHNLTRFOPTALYRLPLVYTPPELPGRAKLAIVLTGVLFALAL 60

QY 61 FGNALVYVYVTRSKAMRTVTNIFICSLALSLLITFFCIPVTMLQNSIDNWLGAFICCKM 120

DB 61 FGNALVYVYVTRSKAMRTVTNIFICSLALSLLITFFCIPVTMLQNSIDNWLGAFICCKM 120



QY 121 VPFVOSTAVVT[ILMTTCIAVERHOGVLVHPFKMKWQVNTNRRFTMLGCVVNLVAVIVGSPM 180  
 DB 121 VPFVOSTAVVT[ILMTTCIAVERHOGVLVHPFKMKWQVNTNRRFTMLGCVVNLVAVIVGSPM 180  
 QY 181 WHVOOLEIKYDFLYPKEHICCLEFWTSVHOKIYVTTFLVILFLPLMVMILLYSKIGYE 240  
 DB 181 WHVOOLEIKYDFLYPKEHICCLEFWTSVHOKIYVTTFLVILFLPLMVMILLYSKIGYE 240  
 QY 241 LWIKKRVGSGVLRITIHGKEMSKTARKKRAVINMVTWVAFVACWAPFHVHMMIEYSN 300  
 DB 241 LWIKKRVGSGVLRITIHGKEMSKTARKKRAVINMVTWVAFVACWAPFHVHMMIEYSN 300  
 QY 301 FEKEYDDVTIRKMFALVQIIGFSNSICNPVYAFPMNENFKKNVLSAVCYCIVNKTSPAQ 360  
 DB 301 FEKEYDDVTIRKMFALVQIIGFSNSICNPVYAFPMNENFKKNVLSAVCYCIVNKTSPAQ 360  
 QY 361 RHGNSGITMMRKKAFFSLRENPNVEETKGEAFSDGNI EVKICQTEEEKKLKRHLAIFRSE 420  
 DB 361 RHGNSGITMMRKKAFFSLRENPNVEETKGEAFSDGNI EVKICQTEEEKKLKRHLAIFRSE 420  
 QY 421 LAENSPDLSGH 431  
 DB 421 LAENSPDLSGH 431

## RESULT 6

AAU76035

ID AAU76035 standard; Protein, 431 AA.

XX AAU76035;

AC AAU76035;

XX AAU76035;

DB AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

XX AAU76035;

CC that binds to NPY-like GPCR can be used to detect it. The NPY-like GPCR  
 CC can be used to screen for agents that can regulate the activity of it. A  
 CC reagent that binds to or regulates NPY-like GPCR can be used to reduce  
 CC the biological activity of NPY-like GPCR, or treat a disorder associated  
 CC with abnormal expression of NPY-like GPCR. The reagent is useful for  
 CC preventing, ameliorating or correcting dysfunctions or diseases such as  
 CC obesity, anorexia, bulimia, diabetes, hypotension, hypertension, cocaine  
 CC withdrawal, pain, cancers, asthma, urinary retention, osteoporosis,  
 CC angina pectoris, myocardial infarction, ulcers, allergies, congestive  
 CC heart failure, memory enhancement, cardiac and cerebral vasospasm,  
 CC pheochromocytoma, ganglioneuroblastoma, Huntington's disease, severe  
 CC mental retardation, Alzheimer's disease, Parkinson's disease, psychotic  
 CC and neurological disorders including anxiety, schizophrenia, manic  
 CC depression, delirium, dementia and dyskinesias. The polypeptide may also  
 CC be used in raising specific antibodies. The present amino acid sequence  
 CC represents human NPY-like GPCR protein of the invention. This protein is  
 CC encoded by the human NPY-like GPCR gene located on chromosome 11q12.2.  
 XX  
 SQ Sequence 431 AA;

Query Match

Best Local Similarity 100.0%; Score 2235; DB 23; Length 431;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQFSRLRDHNTREOFIALYRLRFLVYTPFLPGRAKIALVLTVLIFAL 60  
 DB 1 MQALNITPEQFSRLRDHNTREOFIALYRLRFLVYTPFLPGRAKIALVLTVLIFAL 60

QY 61 FGNALVYVVTIRSKAMRTVNTIFCSIALSDLLITFFCIPTVMTLQNSDNLGAFICM 120  
 DB 61 FGNALVYVVTIRSKAMRTVNTIFCSIALSDLLITFFCIPTVMTLQNSDNLGAFICM 120

QY 121 VPFVOSTAVVT[ILMTTCIAVERHOGVLVHPFKMKWQVNTNRRFTMLGCVVNLVAVIVGSPM 180  
 DB 121 VPFVOSTAVVT[ILMTTCIAVERHOGVLVHPFKMKWQVNTNRRFTMLGCVVNLVAVIVGSPM 180

QY 181 WHVOOLEIKYDFLYPKEHICCLEFWTSVHOKIYVTTFLVILFLPLMVMILLYSKIGYE 240  
 DB 181 WHVOOLEIKYDFLYPKEHICCLEFWTSVHOKIYVTTFLVILFLPLMVMILLYSKIGYE 240

QY 241 LWIKKRVGSGVLRITIHGKEMSKTARKKRAVINMVTWVAFVACWAPFHVHMMIEYSN 300  
 DB 241 LWIKKRVGSGVLRITIHGKEMSKTARKKRAVINMVTWVAFVACWAPFHVHMMIEYSN 300

QY 301 FEKEYDDVTIRKMFALVQIIGFSNSICNPVYAFPMNENFKKNVLSAVCYCIVNKTSPAQ 360  
 DB 301 FEKEYDDVTIRKMFALVQIIGFSNSICNPVYAFPMNENFKKNVLSAVCYCIVNKTSPAQ 360

QY 361 RHGNSGITMMRKKAFFSLRENPNVEETKGEAFSDGNI EVKICQTEEEKKLKRHLAIFRSE 420  
 DB 361 RHGNSGITMMRKKAFFSLRENPNVEETKGEAFSDGNI EVKICQTEEEKKLKRHLAIFRSE 420

QY 421 LAENSPDLSGH 431  
 DB 421 LAENSPDLSGH 431

RESULT 7  
 AAU71309  
 ID AAU71309 standard; Protein, 431 AA.

XX AAU71309;

AC AAU71309;

XX AAU71309;

DT 02-NOV-2000 (first entry)

XX Human orphan G protein-coupled receptor hrup4.

XX Human; orphan G protein-coupled receptor; GPCR; hrup4; drug screening;

XX transmembrane receptor; expressed sequence tag; EST; signal cascade.

OS Homo sapiens.

PN WC2000031258-A2.

Human neuropeptide Y-like GPCR (NPY-like GPCR) protein sequence.  
 Human; neuropeptide Y-like G protein coupled; receptor; antibody;  
 NPY-like GPCR; anorectic; antidiabetic; tranquilizer; hypertensive;  
 hypotensive; analgesic; cytotactic; antiallergic; osteopathic;  
 cardiovascular; antitumor; anticonvulsant; antiallergic;  
 neuroleptic; antidiabetic; antiparkinsonian; neuroprotective; gene therapy;  
 antidepressant; antiallergic; diuretic; chromosome 11q12.2.

Homo sapiens.

WO200204518-A2.

17-JAN-2002.

06-JUL-2001; 2001WO-US21276.

06-JUL-2000; 2000US-216523P.

(FARB ) BAYER CORP.

Bloomquist HT, Zhelein L;

WPI; 2002-148166/19.

N-PSDB; ABK14956.

New neuropeptide Y-like G protein-coupled receptor, for detecting

regulators for preventing, ameliorating, or correcting e.g. obesity,

anorexia, diabetes, hypotension, hypertension, pain, cancers, or

neurological disorders

Claim 10; Fig 1; 91pp; English.

The present invention relates to a new polynucleotide encoding a

polypeptide neuropeptide Y-like G protein-coupled receptor (NPY-like

GPCR), comprising a 431 amino acid sequence, given in the specification.

The polynucleotide of the invention is used to produce a NPY-like GPCR

polypeptide. The complement of the polynucleotide can be used to detect

a coding sequence for a NPY-like GPCR. A reagent, preferably an antibody,

XX 02 JUN 2000.  
 XX 13 OCT 1999; 99W0-US23687.  
 XX 20 NOV 1998; 98US-0109213.  
 XX 16 FEB 1999; 99US-0120416.  
 XX 26 FEB 1999; 99US-0121852.  
 XX 12 MAR 1999; 99US-0123946.  
 XX 12 MAR 1999; 99US-0123949.  
 XX 28 MAY 1999; 99US-0116436.  
 XX 28 MAY 1999; 99US-0116437.  
 XX 28 MAY 1999; 99US-0116439.  
 XX 28 MAY 1999; 99US-0136567.  
 XX 28 MAY 1999; 99US-0137127.  
 XX 28 MAY 1999; 99US-0147131.  
 XX 29 JUN 1999; 99US-0141448.  
 XX 29 SEP 1999; 99US-0156555.  
 XX 29 SEP 1999; 99US-0156633.  
 XX 29 SEP 1999; 99US-0156634.  
 XX 01 OCT 1999; 99US-0157280.  
 XX 01 OCT 1999; 99US-0157281.  
 XX 01 OCT 1999; 99US-0157282.  
 XX 01 OCT 1999; 99US-0157293.  
 XX 01 OCT 1999; 99US-0157294.  
 XX 12 OCT 1999; 99US-0416760.  
 XX 12 OCT 1999; 99US-0417044.  
 XX (AREN-) ARENA PHARM INC.  
 XX Chen R, Dang HT, Liaw CW, Lin I;  
 XX WPI: 2000-40068/34  
 XX N-PSDB: AAD01116.  
 XX Novel human orphan G protein-coupled receptors and the encoding cDNAs  
 XX for use in the identification of G protein-coupled receptor agonists -  
 XX Claim 74; Page 89-91; 107pp; English  
 XX The present amino acid sequence is the hRUP4, an endogenous human  
 XX orphan G protein-coupled receptor (GPCR). The full length hRUP4 cDNA was  
 XX cloned by RT-PCR with human brain cDNA as template. The hRUP4 PCR  
 XX fragment obtained was an alternatively spliced form of the EST (expressed  
 XX sequence tag) clone A107658. The orphan GPCR of the invention, like  
 XX all GPCRs has seven transmembrane alpha helices with an extracellular  
 XX N-terminus and an intracellular C-terminus. However, no endogenous  
 XX ligands has yet been identified for the proteins of the invention. The  
 XX orphan GPCRs may be used in the identification of their endogenous  
 XX ligands, and to screen potential GPCR agonists and antagonists for use as  
 XX pharmaceutical agents. The proteins may also be used in the study of  
 XX GPCR-mediated signalling cascades, and to elucidate their precise role in  
 XX normal and diseased human conditions. Nucleic acid encoding human orphan  
 XX GPCRs may be used for tissue localisation expression analysis to provide  
 XX information about their function in healthy and pathological states.  
 XX Sequence 441 AA;  
 XX  
 XX Query Match 99.9%; Score 2232; DB 21; Length 431;  
 XX Best Local Similarity 99.8%; Pred. No. 6, 1e-237;  
 XX Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 XX 1 MQALNITPQFSLKDHNLTRQFIALYRLRPLVYTPELPGRAKIALVLTGLTALAL 60  
 XX 1 MQALNITPQFSLKDHNLTRQFIALYRLRPLVYTPELPGRAKIALVLTGLTALAL 60  
 XX 61 FGNALVYVYVTRSKAKRTVNTIFCSLSDLLITFFCIPVTMLQNSDNLGGAFICKM 120  
 XX 61 FGNALVYVYVTRSKAKRTVNTIFCSLSDLLITFFCIPVTMLQNSDNLGGAFICKM 120  
 XX 121 VPFVQSTAVVTEMLTMTCTAVERHQGLVHPFKMKWLYTNKRAF-TMLGVVWLVAVIVGSPM 180  
 XX 121 VPFVQSTAVVTEMLTMTCTAVERHQGLVHPFKMKWLYTNKRAF-TMLGVVWLVAVIVGSPM 180

DB 121 VPFVQSTAVVTEMLTMTCTAVERHQGLVHPFKMKWLYTNKRAF-TMLGVVWLVAVIVGSPM 180  
 QY 181 WHVOOLEIKYDFLYEKHEICCLFEWISPVHOKIYITFFILVILFLLPLMVMILLYSKICYE 240  
 DB 181 WHVOOLEIKYDFLYEKHEICCLFEWISPVHOKIYITFFILVILFLLPLMVMILLYSKICYE 240  
 QY 241 LWIKKPVGDSGLFTIHRKEMSKIAKKKKPAVIMMVTVALFAVWAPFHVHMMIEYSN 300  
 DB 241 LWIKKRVGDSGLRTIHGKEMSKIAKKKKRAVIMMVTVALFAVCAWPFHVHMMIEYSN 300  
 QY 301 FEKEYDDVTIKMIFAVQIIGFSNSICNDIVYAFNNENFKKVI SAVCYCIVNKTFSQAQ 360  
 DB 301 FEKEYDDVTIKMIFAVQIIGFSNSICNDIVYAFNNENFKKVI SAVCYCIVNKTFSQAQ 360  
 QY 361 RHGNSGITMMRKKAKFSLRNPVETKGFASDGNIEVKLCHQTEKKKIAKHIALPSE 420  
 DB 361 RHGNSGITMMRKKAKFSLRNPVETKGFASDGNIEVKLCHQTEKKKIAKHIALPSE 420  
 QY 421 LAENSPDLSGH 431  
 DB 421 LAENSPDLSGH 431  
 RESULT 8  
 AAB02843  
 ID AAB02843 standard; Protein, 431 AA.  
 XX AAB02843;  
 XX 22-AUG-2000 (first entry)  
 DE Human G protein coupled receptor hRUP4 protein SEQ ID NO:40.  
 DE Human; G protein coupled receptor, GPCR; transmembrane receptor;  
 KW identification; agonist, screening, therapeutic; pharmaceutical;  
 KW mutant.  
 XX Homo sapiens.  
 XX WO2000022131-A2.  
 XX 20-APR-2000.  
 PD 13-OCT-1999; 99W0-US24065.  
 PF 13-OCT-1999; 98US-0170496.  
 XX 13-OCT-1998; 98US-0108029.  
 PR 12-NOV-1998; 98US-0109213.  
 PR 20-NOV-1998; 98US-0110060.  
 PR 27-NOV-1998; 98US-0120416.  
 PR 16-FEB-1999; 99US-0121852.  
 PR 26-FEB-1999; 99US-0123944.  
 PR 12-MAR-1999; 99US-0123945.  
 PR 12-MAR-1999; 99US-0123946.  
 PR 12-MAR-1999; 99US-0123948.  
 PR 12-MAR-1999; 99US-0123949.  
 PR 12-MAR-1999; 99US-0123951.  
 PR 28-MAR-1999; 99US-0136436.  
 PR 28-MAR-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 28-MAY-1999; 99US-0137567.  
 PR 30-JUN-1999; 99US-0141448.  
 PR 27-AUG-1999; 99US-0151114.  
 PR 03-SEP-1999; 99US-0152524.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156655.  
 PR 29-SEP-1999; 99US-0156634.  
 XX (AREN-) ARENA PHARM INC.  
 XX Behan DP, Letmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;

PI Core M, Liaw CW, Lin I, Lowitz K, White C;  
 XX WPI: 2000-317986/27.  
 DR N-PSDB; AAA46037  
 XX  
 PT Non-endogenous, human G protein-coupled receptors for screening  
 PT receptor, inverse or partial agonists useful as therapeutic agents  
 XX  
 PS Example 1: Page 119-120; 187pp; English.  
 XX  
 CC The present invention describes transmembrane receptors, preferably  
 CC human G protein coupled receptors (GPCR), for which the endogenous  
 CC ligand is unknown (orphan GPCR receptors). More specifically the present  
 CC invention relates to non-endogenous, constitutively activated versions  
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
 CC the direct identification of candidate compounds as receptors agonists,  
 CC inverse agonists or partial agonists for use as pharmaceutical agents.  
 CC AAA46017 to AAA46126 and AAA02825 to AAA02859 represent sequences used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 431 AA:  
 Query Match 99.9%; Score 2232, DB 21, Length 431.  
 Best Local Similarity 99.8%; Pred. No. 6.1e-237;  
 Matches 430, Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MQALNITPEQSRLLRDHNTPEQFIALYRLPLVLYTPPELPGAKLALVLTGLIFALAL 60  
 DB 1 MQALNITPEQSRLLRDHNTPEQFIALYRLPLVLYTPPELPGAKLALVLTGLIFALAL 60  
 QY 61 FGNALVFYVYVTRSKAMRTVTNIFTCSLALSDLLITFCIPVTMLQNI SDNWLGAFICKM 120  
 DB 61 FGNALVFYVYVTRSKAMRTVTNIFTCSLALSDLLITFCIPVTMLQNI SDNWLGAFICKM 120  
 QY 121 VPFGOSTAVVTETILTMTCTIAVERHGLVHPFKMKQYTNKSKAFITMLGVWLVAVIVGSPM 180  
 DB 121 VPFGOSTAVVTETILTMTCTIAVERHGLVHPFKMKQYTNKSKAFITMLGVWLVAVIVGSPM 180  
 QY 181 MHVQOLEIKYDFLYEKEHICCLEWTSVPVHQKIYTTFTILVILFLLPLMVLILYSKIGYE 240  
 DB 181 MHVQOLEIKYDFLYEKEHICCLEWTSVPVHQKIYTTFTILVILFLLPLMVLILYSKIGYE 240  
 QY 241 LWIKRRVGDGSLRTIHGKEMSKLARKKKRAVIMVTVVWALFVACWAPFHVHMMIEYSN 300  
 DB 241 LWIKRRVGDGSLRTIHGKEMSKLARKKKRAVIMVTVVWALFVACWAPFHVHMMIEYSN 300  
 QY 301 FEKEYDDVTIKMIFAVOIIIGFSNSICNPIVYAFMNEFKKNVLSAVCYCIVNKTSPAQ 360  
 DB 301 FEKEYDDVTIKMIFAVOIIIGFSNSICNPIVYAFMNEFKKNVLSAVCYCIVNKTSPAQ 360  
 QY 361 RHGNSGITMRRKKAKFSIRENPVEETKGEAFSDGNIIEVKLCEQTEERKKLRLHALFRSE 420  
 DB 361 RHGNSGITMRRKKAKFSIRENPVEETKGEAFSDGNIIEVKLCEQTEERKKLRLHALFRSE 420  
 QY 421 LAENSPIDSGH 431  
 DB 421 LAENSPIDSGH 431

## RESULT 9

AA02853  
 ID AA02853 standard, Protein; 431 AA.

XX  
 AC AA02853;

DT 22-AUG-2000 (first entry)

XX Human G protein coupled receptor hRUF4 (V272K) protein SEQ ID NO:128.

DE Human: G protein coupled receptor; GPCR: transmembrane receptor,  
 KW identification; agonist; screening; therapeutic; pharmaceutical;  
 KW mutant.

XX

OS Homo sapiens.  
 OS Synthetic.  
 XX WC200002141-A2.  
 XX 20-APR-2000.  
 XX  
 XX 13-OCT-1999; 99WO-US24065.  
 XX  
 PR 13-OCT-1998; 98US-0170496.  
 PR 12-NOV-1998; 98US-0108029.  
 PR 20-NOV-1998; 98US-0109213.  
 PR 27-NOV-1998; 98US-0110060.  
 PR 16-FEB-1999; 99US-0120416.  
 PR 26-FEB-1999; 99US-0121852.  
 PR 12-MAR-1999; 99US-0123944.  
 PR 12-MAR-1999; 99US-0123945.  
 PR 12-MAR-1999; 99US-0123946.  
 PR 12-MAR-1999; 99US-0123948.  
 PR 12-MAR-1999; 99US-0123949.  
 PR 12-MAR-1999; 99US-0123951.  
 PR 28-MAY-1999; 99US-0136436.  
 PR 28-MAY-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0136439.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 30-JUN-1999; 99US-0137567.  
 PR 27-AUG-1999; 99US-0141448.  
 PR 03-SEP-1999; 99US-0152524.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156555.  
 PR 29-SEP-1999; 99US-0156634.  
 XX  
 PA (AREN-) ARENA PHARM INC.

Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;  
 Gore M, Liaw CW, Lin I, Lowitz K, White C;

WPI: 2000-317986/27.  
 N-PSDB; AAA46115.

Non-endogenous, human G protein-coupled receptors for screening  
 receptor, inverse or partial agonists useful as therapeutic agents

Example 2: Page 164-166; 187pp; English.

The present invention describes transmembrane receptors, preferably  
 human G protein coupled receptors (GPCR), for which the endogenous  
 ligand is unknown (orphan GPCR receptors). More specifically the present  
 invention relates to non-endogenous, constitutively activated versions  
 of a human GPCR. These non-endogenous human GPCRs can be useful for  
 the direct identification of candidate compounds as receptors agonists,  
 inverse agonists or partial agonists for use as pharmaceutical agents.  
 CC AAA46017 to AAA46126 and AAA02825 to AAA02859 represent sequences used in  
 CC the exemplification of the present invention.

XX Sequence 431 AA:

Query Match 99.6%; Score 2226; DB 21; Length 431;  
 Best Local Similarity 99.5%; Pred. No. 2.8e-236;  
 Matches 429; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MQALNITPEQSRLLRDHNTPEQFIALYRLPLVLYTPPELPGAKLALVLTGLIFALAL 60  
 DB 1 MQALNITPEQSRLLRDHNTPEQFIALYRLPLVLYTPPELPGAKLALVLTGLIFALAL 60  
 QY 61 FGNALVFYVYVTRSKAMRTVTNIFTCSLALSDLLITFCIPVTMLQNI SDNWLGAFICKM 120  
 DB 61 FGNALVFYVYVTRSKAMRTVTNIFTCSLALSDLLITFCIPVTMLQNI SDNWLGAFICKM 120  
 QY 121 VPFGOSTAVVTETILTMTCTIAVERHGLVHPFKMKQYTNKSKAFITMLGVWLVAVIVGSPM 180  
 XX

121 VPEVQSTAVVTEMLTMTCTIAVERHQGLVHPFKMKWGYTNRRAFTMLGVVWVAVIVGSPM 180  
 181 WHVQOLEIKYDFLYEREHICCLEWTSVHQKIYTFILVILFLLPLVMMLILYSKIGYE 240  
 181 WHVQOLEIKYDFLYEREHICCLEWTSVHQKIYTFILVILFLLPLVMMLILYSKIGYE 240  
 241 LMKKKVQDGSVLTIRGKEMSKIAKKKRAVIMVTVVAVFACWAPFHVHMMIYYSN 300  
 241 LMKKKVQDGSVLTIRGKEMSKIAKKKRAVIMVTVVAVFACWAPFHVHMMIYYSN 300  
 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFNENFKKVLNAVYCIVNKTFSQAQ 360  
 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFNENFKKVLNAVYCIVNKTFSQAQ 360  
 361 RHGNSGITMMKKAKESLPFNPFETKIEAFSGKNTIFVKI/ETETKKKKLPHLALPSE 420  
 361 RHGNSGITMMKKAKESLPFNPFETKIEAFSGKNTIFVKI/ETETKKKKLPHLALPSE 420  
 421 LAENSPDLSGH 431  
 421 LAENSPDLSGH 431

## RESULT 10

AAU97740

AAU97740 standard; Protein: 433 AA.

XX AAU97740;

XX 27 AUG 2002 (first entry)

XX Mouse TGR346a polypeptide.

XX Mouse: TGR346a: receptor; G-protein coupled receptor; GPCR; TGR.  
 KW TGR-associated disorder; signal transduction; renal failure; nephritis;  
 KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;  
 KW diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;  
 KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;  
 KW circadian rhythm.

XX Mus sp.

XX W020024245H A2.

XX 30 MAY 2002.

XX 21 NOV 2001; 2001W0-0543404.

XX 22 NOV 2000; 2000US 252841P.

XX 22 DEC 2000; 2000US 257646P.

XX 12-JAN-2001; 2001US 261377P.

XX 28-MAR-2001; 2001US-279554P.

XX 29-MAR 2001; 2001US 280676P.

XX (TULA ) TULARIK INC.

XX Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;

XX WPI; 2002 464644/49.

XX N PDSH; AHR86292.

XX New isolated G-protein coupled receptor polypeptide, termed TGR, for  
 PT diagnosis and treatment of diseases such as renal failure, nephritis,  
 PT hypothyroidism, diabetes insipidus, and disturbances of thirst and  
 PT sleep.

XX Claim 29; Page 78; 98pp; English.

XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),  
 CC termed TGR, and its associated nucleic acid. The sequences of the  
 CC invention are useful for identifying a compound that modulates signal  
 CC transduction and for identifying a mammal having a TGR-associated  
 CC disorder. The proteins and nucleic acids are useful in diagnosis and

CC treatment of diseases or conditions such as renal failure, nephritis,  
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,  
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,  
 CC sleep, temperature regulation, appetite, blood pressure or any other  
 CC syndrome or disease associated with the hypothalamus. The sequences can  
 CC be used in regulation of circadian rhythms, for use as genetic markers  
 CC for the identification of mutations associated with diseases resulting  
 CC from GPCR inactivation in particular cell types and for identification of  
 CC modulators of GPCR signal transduction. This sequence represents the  
 CC mouse TGR346a polypeptide.

XX Sequence 433 AA;

SQ Query Match 85.6%; Score 1914; DB 23; Length 433;

Best Local Similarity 83.5%; Pred. No. 6.9e 202;

Matches 360; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

QY 1 MQALNITPPQFSKILRDHNLITREQFIALYRIURPIVYTPELPCRALKALIVTCVLIIFALAL 60  
 DB 1 MQALNITAPQFSKILSAHNLITREQFIHRYGLRPLVYTPELPAKALAFALAGALIFALAL 60  
 QY 61 FGNALVYVYVTPSKAMPTVTNITFICSLALSLLITFFCIPVTMLQNSDNLGGAFICKM 120  
 DB 61 FGNLSLVIYVTRSKAMRTVTNITFICSLALSLLITFFCIPVTMLQNSDNLGGAFICKM 120  
 QY 121 VPEVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKWGYTNRRAFTMLGVVWVAVIVGSPM 180  
 DB 121 VPEVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKWGYTNRRAFTMLGVVWVAVIVGSPM 180  
 QY 181 WHVQOLEIKYDFLYEREHICCLEWTSVHQKIYTFILVILFLLPLVMMLILYSKIGYE 240  
 DB 181 WHVQOLEIKYDFLYEREHICCLEWTSVHQKIYTFILVILFLLPLVMMLILYSKIGYE 240  
 QY 241 LMKKKVQDGSVLTIRGKEMSKIAKKKRAVIMVTVVAVFACWAPFHVHMMIYYSN 300  
 DB 241 LMKKKVQDGSVLTIRGKEMSKIAKKKRAVIMVTVVAVFACWAPFHVHMMIYYSN 300  
 QY 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFNENFKKVLNAVYCIVNKTFSQAQ 360  
 DB 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFNENFKKVLNAVYCIVNKTFSQAQ 360  
 QY 361 RHGNSGITMMKKAKESLPFNPFETKIEAFSGKNTIFVKI/ETETKKKKLPHLALPSE 420  
 DB 361 RHGNSGITMMKKAKESLPFNPFETKIEAFSGKNTIFVKI/ETETKKKKLPHLALPSE 420  
 QY 421 LAENSPDLSGH 431  
 DB 421 LAENSPDLSGH 431

## RESULT 11

AAU97741

AAU97741 standard; Protein: 416 AA.

XX AAU97741;

XX 27-AUG-2002 (first entry)

XX Mouse TGR346b polypeptide.

XX Mouse: TGR346b: receptor; G-protein coupled receptor; GPCR; TGR;  
 KW TGR-associated disorder; signal transduction; renal failure; nephritis;  
 KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;  
 KW diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;  
 KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;  
 KW circadian rhythm.

XX Mus sp.

XX W020024245H A2.

XX 30-MAY-2002.

XX



Homosapiens.

PT  
PT

XX  
PS Claim 11. Page 8-9. 2700. English

[illegible]

CC This sequence represents the human

CC The protein can be used to isolate agonists and antagonists.

CC be used as active agents in the treatment of infections (e.g.

fungus, protozoal and viral infections, particularly HIV-1 o

CC pain, cancers, anorexia, bulimia, asthma, Parkinson's disease  
CC heart failure, hypertension, heart disease, major depression

CC heart failure, hypotension, urinary retention, urinary retention, ulcer  
CC osteoporosis, angina pectoris, myocardial infarction, ulcers

allergics, benign prostatic hypertrophy and psychotic and neurotic disorders, and, as peccators, myocardial infarction, griefs

CC disorders.

XX

| SQ  | Sequence | 420 AA; |
|-----|----------|---------|
| 1   | 1        | 1       |
| 2   | 2        | 2       |
| 3   | 3        | 3       |
| 4   | 4        | 4       |
| 5   | 5        | 5       |
| 6   | 6        | 6       |
| 7   | 7        | 7       |
| 8   | 8        | 8       |
| 9   | 9        | 9       |
| 10  | 10       | 10      |
| 11  | 11       | 11      |
| 12  | 12       | 12      |
| 13  | 13       | 13      |
| 14  | 14       | 14      |
| 15  | 15       | 15      |
| 16  | 16       | 16      |
| 17  | 17       | 17      |
| 18  | 18       | 18      |
| 19  | 19       | 19      |
| 20  | 20       | 20      |
| 21  | 21       | 21      |
| 22  | 22       | 22      |
| 23  | 23       | 23      |
| 24  | 24       | 24      |
| 25  | 25       | 25      |
| 26  | 26       | 26      |
| 27  | 27       | 27      |
| 28  | 28       | 28      |
| 29  | 29       | 29      |
| 30  | 30       | 30      |
| 31  | 31       | 31      |
| 32  | 32       | 32      |
| 33  | 33       | 33      |
| 34  | 34       | 34      |
| 35  | 35       | 35      |
| 36  | 36       | 36      |
| 37  | 37       | 37      |
| 38  | 38       | 38      |
| 39  | 39       | 39      |
| 40  | 40       | 40      |
| 41  | 41       | 41      |
| 42  | 42       | 42      |
| 43  | 43       | 43      |
| 44  | 44       | 44      |
| 45  | 45       | 45      |
| 46  | 46       | 46      |
| 47  | 47       | 47      |
| 48  | 48       | 48      |
| 49  | 49       | 49      |
| 50  | 50       | 50      |
| 51  | 51       | 51      |
| 52  | 52       | 52      |
| 53  | 53       | 53      |
| 54  | 54       | 54      |
| 55  | 55       | 55      |
| 56  | 56       | 56      |
| 57  | 57       | 57      |
| 58  | 58       | 58      |
| 59  | 59       | 59      |
| 60  | 60       | 60      |
| 61  | 61       | 61      |
| 62  | 62       | 62      |
| 63  | 63       | 63      |
| 64  | 64       | 64      |
| 65  | 65       | 65      |
| 66  | 66       | 66      |
| 67  | 67       | 67      |
| 68  | 68       | 68      |
| 69  | 69       | 69      |
| 70  | 70       | 70      |
| 71  | 71       | 71      |
| 72  | 72       | 72      |
| 73  | 73       | 73      |
| 74  | 74       | 74      |
| 75  | 75       | 75      |
| 76  | 76       | 76      |
| 77  | 77       | 77      |
| 78  | 78       | 78      |
| 79  | 79       | 79      |
| 80  | 80       | 80      |
| 81  | 81       | 81      |
| 82  | 82       | 82      |
| 83  | 83       | 83      |
| 84  | 84       | 84      |
| 85  | 85       | 85      |
| 86  | 86       | 86      |
| 87  | 87       | 87      |
| 88  | 88       | 88      |
| 89  | 89       | 89      |
| 90  | 90       | 90      |
| 91  | 91       | 91      |
| 92  | 92       | 92      |
| 93  | 93       | 93      |
| 94  | 94       | 94      |
| 95  | 95       | 95      |
| 96  | 96       | 96      |
| 97  | 97       | 97      |
| 98  | 98       | 98      |
| 99  | 99       | 99      |
| 100 | 100      | 100     |

[illegible]

|                       |                                      |
|-----------------------|--------------------------------------|
| Query Match           | 22.58; Score 502; DH 20; Length 420; |
| Host Local Similarity | 30.18; Prod No. 3 000-46.            |

best local similarity 30.1%, Pled: NO. 2.90-46;  
Matches 128: Conservative 87: Mismatches 162: Indels 4

[illegible]

QY 14 LLRDHNLTRQFIALYRLRPLVYTPELPGRAKLALVLTGVLIFALALFGNA

1.  $\frac{1}{2}$

2.  $\frac{1}{2}$

3.  $\frac{1}{2}$

4.  $\frac{1}{2}$

5.  $\frac{1}{2}$

6.  $\frac{1}{2}$

7.  $\frac{1}{2}$

8.  $\frac{1}{2}$

9.  $\frac{1}{2}$

10.  $\frac{1}{2}$

11.  $\frac{1}{2}$

12.  $\frac{1}{2}$

13.  $\frac{1}{2}$

14.  $\frac{1}{2}$

15.  $\frac{1}{2}$

16.  $\frac{1}{2}$

17.  $\frac{1}{2}$

18.  $\frac{1}{2}$

19.  $\frac{1}{2}$

20.  $\frac{1}{2}$

21.  $\frac{1}{2}$

22.  $\frac{1}{2}$

23.  $\frac{1}{2}$

24.  $\frac{1}{2}$

25.  $\frac{1}{2}$

26.  $\frac{1}{2}$

27.  $\frac{1}{2}$

28.  $\frac{1}{2}$

29.  $\frac{1}{2}$

30.  $\frac{1}{2}$

31.  $\frac{1}{2}$

32.  $\frac{1}{2}$

33.  $\frac{1}{2}$

34.  $\frac{1}{2}$

35.  $\frac{1}{2}$

36.  $\frac{1}{2}$

37.  $\frac{1}{2}$

38.  $\frac{1}{2}$

39.  $\frac{1}{2}$

40.  $\frac{1}{2}$

41.  $\frac{1}{2}$

42.  $\frac{1}{2}$

43.  $\frac{1}{2}$

44.  $\frac{1}{2}$

45.  $\frac{1}{2}$

46.  $\frac{1}{2}$

47.  $\frac{1}{2}$

48.  $\frac{1}{2}$

49.  $\frac{1}{2}$

50.  $\frac{1}{2}$

51.  $\frac{1}{2}$

52.  $\frac{1}{2}$

53.  $\frac{1}{2}$

54.  $\frac{1}{2}$

55.  $\frac{1}{2}$

56.  $\frac{1}{2}$

57.  $\frac{1}{2}$

58.  $\frac{1}{2}$

59.  $\frac{1}{2}$

60.  $\frac{1}{2}$

61.  $\frac{1}{2}$

62.  $\frac{1}{2}$

63.  $\frac{1}{2}$

64.  $\frac{1}{2}$

65.  $\frac{1}{2}$

66.  $\frac{1}{2}$

67.  $\frac{1}{2}$

68.  $\frac{1}{2}$

69.  $\frac{1}{2}$

70.  $\frac{1}{2}$

71.  $\frac{1}{2}$

72.  $\frac{1}{2}$

73.  $\frac{1}{2}$

74.  $\frac{1}{2}$

75.  $\frac{1}{2}$

76.  $\frac{1}{2}$

77.  $\frac{1}{2}$

78.  $\frac{1}{2}$

79.  $\frac{1}{2}$

80.  $\frac{1}{2}$

81.  $\frac{1}{2}$

82.  $\frac{1}{2}$

83.  $\frac{1}{2}$

84.  $\frac{1}{2}$

85.  $\frac{1}{2}$

86.  $\frac{1}{2}$

87.  $\frac{1}{2}$

88.  $\frac{1}{2}$

89.  $\frac{1}{2}$

90.  $\frac{1}{2}$

91.  $\frac{1}{2}$

92.  $\frac{1}{2}$

93.  $\frac{1}{2}$

94.  $\frac{1}{2}$

95.  $\frac{1}{2}$

96.  $\frac{1}{2}$

97.  $\frac{1}{2}$

98.  $\frac{1}{2}$

99.  $\frac{1}{2}$

100.  $\frac{1}{2}$

26 I,YSDINI'T---YVNYYL,HQPQV-----AAIF|ISYFI,IFFI,CMMGNT

74 KAMPOTTEVNILICCI KICCI I IUDIC FUDIMI ONI CIOBBI CACU CUDAGI...

QY 74 KAMRIVNIFCSIALSD/LIIFFCIPVIMQNISDNWIGGAFICKMVPFV

74 KHMITVTNLFILNLAISDLLVGFECMPITLIDNITAGWPFGNTMCKISGLV

1. **THEORY OF THE CASE**

134 1.TMTCTC1A V F R H Q G L V H P F K M K W Q Y T N R R A F T M I G V V W I V A V I V G S P - - M W Q Y

134 FTI.VAIAVDRFQC VVYPFKPK - - I'I'IK'I'A'F'V I I M I I WVI.A I'I'IMS pSAVMI,

101 DER VEREIN - - - - - 1001 POLITISCHER VEREIN DER VEREIN DER VEREIN

QY I 31 DECI ERM - - - - - IC C DEEWISFVHQK IY I I F I LV I L F L L P L M V M L I L Y S K

191 RVRLNSONKTSPPVYWCREDWPNOEMRK YTTTVLFANVYAPISLIVIMYCR

**PROCEEDINGS OF THE**

246 RVGIDGSVLRTIHGKEMSK-----[ARKKKRAVIMMVTVVALF'AVCWAIP'FHVV

1.  $\frac{1}{2}$  2.  $\frac{1}{3}$  3.  $\frac{1}{4}$  4.  $\frac{1}{5}$  5.  $\frac{1}{6}$  6.  $\frac{1}{7}$  7.  $\frac{1}{8}$  8.  $\frac{1}{9}$  9.  $\frac{1}{10}$  10.  $\frac{1}{11}$  11.  $\frac{1}{12}$  12.  $\frac{1}{13}$  13.  $\frac{1}{14}$  14.  $\frac{1}{15}$  15.  $\frac{1}{16}$  16.  $\frac{1}{17}$  17.  $\frac{1}{18}$  18.  $\frac{1}{19}$  19.  $\frac{1}{20}$  20.  $\frac{1}{21}$  21.  $\frac{1}{22}$  22.  $\frac{1}{23}$  23.  $\frac{1}{24}$  24.  $\frac{1}{25}$  25.  $\frac{1}{26}$  26.  $\frac{1}{27}$  27.  $\frac{1}{28}$  28.  $\frac{1}{29}$  29.  $\frac{1}{30}$  30.  $\frac{1}{31}$  31.  $\frac{1}{32}$  32.  $\frac{1}{33}$  33.  $\frac{1}{34}$  34.  $\frac{1}{35}$  35.  $\frac{1}{36}$  36.  $\frac{1}{37}$  37.  $\frac{1}{38}$  38.  $\frac{1}{39}$  39.  $\frac{1}{40}$  40.  $\frac{1}{41}$  41.  $\frac{1}{42}$  42.  $\frac{1}{43}$  43.  $\frac{1}{44}$  44.  $\frac{1}{45}$  45.  $\frac{1}{46}$  46.  $\frac{1}{47}$  47.  $\frac{1}{48}$  48.  $\frac{1}{49}$  49.  $\frac{1}{50}$  50.  $\frac{1}{51}$  51.  $\frac{1}{52}$  52.  $\frac{1}{53}$  53.  $\frac{1}{54}$  54.  $\frac{1}{55}$  55.  $\frac{1}{56}$  56.  $\frac{1}{57}$  57.  $\frac{1}{58}$  58.  $\frac{1}{59}$  59.  $\frac{1}{60}$  60.  $\frac{1}{61}$  61.  $\frac{1}{62}$  62.  $\frac{1}{63}$  63.  $\frac{1}{64}$  64.  $\frac{1}{65}$  65.  $\frac{1}{66}$  66.  $\frac{1}{67}$  67.  $\frac{1}{68}$  68.  $\frac{1}{69}$  69.  $\frac{1}{70}$  70.  $\frac{1}{71}$  71.  $\frac{1}{72}$  72.  $\frac{1}{73}$  73.  $\frac{1}{74}$  74.  $\frac{1}{75}$  75.  $\frac{1}{76}$  76.  $\frac{1}{77}$  77.  $\frac{1}{78}$  78.  $\frac{1}{79}$  79.  $\frac{1}{80}$  80.  $\frac{1}{81}$  81.  $\frac{1}{82}$  82.  $\frac{1}{83}$  83.  $\frac{1}{84}$  84.  $\frac{1}{85}$  85.  $\frac{1}{86}$  86.  $\frac{1}{87}$  87.  $\frac{1}{88}$  88.  $\frac{1}{89}$  89.  $\frac{1}{90}$  90.  $\frac{1}{91}$  91.  $\frac{1}{92}$  92.  $\frac{1}{93}$  93.  $\frac{1}{94}$  94.  $\frac{1}{95}$  95.  $\frac{1}{96}$  96.  $\frac{1}{97}$  97.  $\frac{1}{98}$  98.  $\frac{1}{99}$  99.  $\frac{1}{100}$  100.  $\frac{1}{101}$  101.  $\frac{1}{102}$  102.  $\frac{1}{103}$  103.  $\frac{1}{104}$  104.  $\frac{1}{105}$  105.  $\frac{1}{106}$  106.  $\frac{1}{107}$  107.  $\frac{1}{108}$  108.  $\frac{1}{109}$  109.  $\frac{1}{110}$  110.  $\frac{1}{111}$  111.  $\frac{1}{112}$  112.  $\frac{1}{113}$  113.  $\frac{1}{114}$  114.  $\frac{1}{115}$  115.  $\frac{1}{116}$  116.  $\frac{1}{117}$  117.  $\frac{1}{118}$  118.  $\frac{1}{119}$  119.  $\frac{1}{120}$  120.  $\frac{1}{121}$  121.  $\frac{1}{122}$  122.  $\frac{1}{123}$  123.  $\frac{1}{124}$  124.  $\frac{1}{125}$  125.  $\frac{1}{126}$  126.  $\frac{1}{127}$  127.  $\frac{1}{128}$  128.  $\frac{1}{129}$  129.  $\frac{1}{130}$  130.  $\frac{1}{131}$  131.  $\frac{1}{132}$  132.  $\frac{1}{133}$  133.  $\frac{1}{134}$  134.  $\frac{1}{135}$  135.  $\frac{1}{136}$  136.  $\frac{1}{137}$  137.  $\frac{1}{138}$  138.  $\frac{1}{139}$  139.  $\frac{1}{140}$  140.  $\frac{1}{141}$  141.  $\frac{1}{142}$  142.  $\frac{1}{143}$  143.  $\frac{1}{144}$  144.  $\frac{1}{145}$  145.  $\frac{1}{146}$  146.  $\frac{1}{147}$  147.  $\frac{1}{148}$  148.  $\frac{1}{149}$  149.  $\frac{1}{150}$  150.  $\frac{1}{151}$  151.  $\frac{1}{152}$  152.  $\frac{1}{153}$  153.  $\frac{1}{154}$  154.  $\frac{1}{155}$  155.  $\frac{1}{156}$  156.  $\frac{1}{157}$  157.  $\frac{1}{158}$  158.  $\frac{1}{159}$  159.  $\frac{1}{160}$  160.  $\frac{1}{161}$  161.  $\frac{1}{162}$  162.  $\frac{1}{163}$  163.  $\frac{1}{164}$  164.  $\frac{1}{165}$  165.  $\frac{1}{166}$  166.  $\frac{1}{167}$  167.  $\frac{1}{168}$  168.  $\frac{1}{169}$  169.  $\frac{1}{170}$  170.  $\frac{1}{171}$  171.  $\frac{1}{172}$  172.  $\frac{1}{173}$  173.  $\frac{1}{174}$  174.  $\frac{1}{175}$  175.  $\frac{1}{176}$  176.  $\frac{1}{177}$  177.  $\frac{1}{178}$  178.  $\frac{1}{179}$  179.  $\frac{1}{180}$  180.  $\frac{1}{181}$  181.  $\frac{1}{182}$  182.  $\frac{1}{183}$  183.  $\frac{1}{184}$  184.  $\frac{1}{185}$  185.  $\frac{1}{186}$  186.  $\frac{1}{187}$  187.  $\frac{1}{188}$  188.  $\frac{1}{189}$  189.  $\frac{1}{190}$  190.  $\frac{1}{191}$  191.  $\frac{1}{192}$  192.  $\frac{1}{193}$  193.  $\frac{1}{194}$  194.  $\frac{1}{195}$  195.  $\frac{1}{196}$  196.  $\frac{1}{197}$  197.  $\frac{1}{198}$  198.  $\frac{1}{199}$  199.  $\frac{1}{200}$  200.  $\frac{1}{201}$  201.  $\frac{1}{202}$  202.  $\frac{1}{203}$  203.  $\frac{1}{204}$  204.  $\frac{1}{205}$  205.  $\frac{1}{206}$  206.  $\frac{1}{207}$  207.  $\frac{1}{208}$  208.  $\frac{1}{209}$  209.  $\frac{1}{210}$  210.  $\frac{1}{211}$  211.  $\frac{1}{212}$  212.  $\frac{1}{213}$  213.  $\frac{1}{214}$  214.  $\frac{1}{215}$  215.  $\frac{1}{216}$  216.  $\frac{1}{217}$  217.  $\frac{1}{218}$  218.  $\frac{1}{219}$  219.  $\frac{1}{220}$  220.  $\frac{1}{221}$  221.  $\frac{1}{222}$  222.  $\frac{1}{223}$  223.  $\frac{1}{224}$  224.  $\frac{1}{225}$  225.  $\frac{1}{226}$  226.  $\frac{1}{227}$  227.  $\frac{1}{228}$  228.  $\frac{1}{229}$  229.  $\frac{1}{230}$  230.  $\frac{1}{231}$  231.  $\frac{1}{232}$  232.  $\frac{1}{233}$  233.  $\frac{1}{234}$  234.  $\frac{1}{235}$  235.  $\frac{1}{236}$  236.  $\frac{1}{237}$  237.  $\frac{1}{238}$  238.  $\frac{1}{239}$  239.  $\frac{1}{240}$  240

248 - - - PAAVPHT - - GPKNQEQWHVVSPPKKQKI KMLLI VALLFILSWLPLWTLIL

\_\_\_\_\_







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:15:43 ; Search time 18 seconds  
(without alignments)  
2301 885 Million cell updates/sec

Title: US-10-070-241B-1

Perfect score: 2235

Sequence: 1 MQAGNITPEQFSPLLDHNL PHLAFRSELAENSPDLSGH 431

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 469.5 | 21.0        | 381    | 2     | 139187      |
| 2          | 446.5 | 20.0        | 375    | 2     | S63F8C      |
| 3          | 426   | 19.1        | 382    | 2     | S27388      |
| 4          | 425.5 | 19.0        | 375    | 2     | G02300      |
| 5          | 425.5 | 19.0        | 375    | 2     | I39182      |
| 6          | 425.5 | 19.0        | 449    | 2     | A41738      |
| 7          | 425   | 19.0        | 382    | 2     | B46133      |
| 8          | 421.5 | 18.9        | 384    | 2     | A45490      |
| 9          | 420.5 | 18.8        | 349    | 2     | S12863      |
| 10         | 420.5 | 18.8        | 349    | 2     | I59336      |
| 11         | 412.5 | 18.5        | 366    | 2     | S71152      |
| 12         | 402.5 | 18.0        | 427    | 2     | S50150      |
| 13         | 401   | 17.9        | 519    | 2     | S17783      |
| 14         | 400.5 | 17.9        | 423    | 2     | B40470      |
| 15         | 398   | 17.8        | 457    | 2     | T29741      |
| 16         | 386.5 | 17.3        | 407    | 2     | S23510      |
| 17         | 384   | 17.2        | 402    | 2     | I56595      |
| 18         | 383.5 | 17.2        | 407    | 1     | J01274      |
| 19         | 382.5 | 17.1        | 444    | 2     | A42685      |
| 20         | 381.5 | 17.1        | 443    | 2     | D40470      |
| 21         | 379   | 17.0        | 428    | 2     | J00692      |
| 22         | 378.5 | 16.9        | 430    | 2     | I51898      |
| 23         | 378   | 16.9        | 385    | 2     | S55524      |
| 24         | 377.5 | 16.9        | 423    | 2     | J02677      |
| 25         | 377.5 | 16.9        | 584    | 2     | J07809      |
| 26         | 375.5 | 16.8        | 412    | 2     | T22076      |
| 27         | 375.5 | 16.8        | 436    | 2     | J05599      |
| 28         | 372   | 16.6        | 452    | 2     | A34916      |
| 29         | 370.5 | 16.6        | 384    | 2     | S20303      |

|    |       |      |     |   |        |
|----|-------|------|-----|---|--------|
| 30 | 370.5 | 16.6 | 399 | 2 | A46632 |
| 31 | 370.5 | 16.6 | 399 | 2 | S29480 |
| 32 | 369   | 16.5 | 384 | 2 | I57957 |
| 33 | 367.5 | 16.4 | 407 | 2 | A34357 |
| 34 | 367.5 | 16.4 | 491 | 2 | C40470 |
| 35 | 366.5 | 16.4 | 407 | 2 | S20304 |
| 36 | 365   | 16.3 | 398 | 1 | J01059 |
| 37 | 364.5 | 16.3 | 384 | 2 | I57682 |
| 38 | 362.5 | 16.2 | 384 | 1 | S00516 |
| 39 | 362.5 | 16.2 | 384 | 2 | A39003 |
| 40 | 362.5 | 16.2 | 465 | 1 | J01517 |
| 41 | 361   | 16.2 | 370 | 1 | I52315 |
| 42 | 360.5 | 16.1 | 390 | 2 | B41007 |
| 43 | 359.5 | 16.1 | 390 | 2 | A36737 |
| 44 | 359   | 16.1 | 384 | 2 | A41007 |
| 45 | 359   | 16.1 | 390 | 2 | JH0374 |

## ALIGNMENTS

### RESULT 1

139187

neuropeptide Y/peptide YY receptor Y2 - human

N:Alternate names: neuropeptide Y/peptide YY receptor type 2

C:Species: Homo sapiens (man)

C:Date: 01-Mar-1996 #sequence revision 01-Mar-1996 #text change 20-Apr-2000

C:Accession: 139187; 139163; G02301

R:Gerald, C.; Walker, M.W.; Vayssie, P.J.

J. Biol. Chem. 270, 26758-26761, 1995

A:Title: Expression cloning and pharmacological characterization of a human hippocamp

A:Reference number: 139187, MUID:960760, PMID:7592910

A:Accession: 139187

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-381 <GER>

A:Cross-references: EMBL:U36269; NID:q1063633; PIDN:AAC50281.1; PID:q1063634

R:Rose, P.M.; Fernandes, P.; Lynch, J.S.; Frazier, S.T.; Fisher, S.M.; Kodukula, K.;

J. Biol. Chem. 270, 22661-22664, 1995

A:Title: Cloning and functional expression of a cDNA encoding a human type 2 neurocep

A:Reference number: 139163; MUID:96032678; PMID:7559383

A:Accession: 139163

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-133; A', 135-381 <ROS>

A:Cross references: EMBL:U3500; NID:q1000750; PIDN:AAA93170.1; PID:q1000751

R:Yan, H.; Yang, J.; Marasco, J.; Yamaguchi, K.; Krenner, S.; Collins, F.; Karbon, W.

submitted to the EMBL Data Library, December 1995

A:Reference number: H01019

A:Accession: G02301

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-171; C', 173-381; E', 201-308; A', 213-381 <YAN>

A:Cross references: EMBL:U34389; NID:q134332; PIDN:AAB07760.1; PID:q134333

C:Genetics.

A:Gene: GDB:NPY2R

A:Cross references: GDB:4365607; OMIM:162642

A:Map position: 4q31-q31

C:Superfamily: neuropeptide Y receptor

F:49-76/Domain: Transmembrane #status predicted <TM>

F:87-113/Domain: Transmembrane #status predicted <TM>

F:166-186/Domain: Transmembrane #status predicted <TM>

F:221-237/Domain: Transmembrane #status predicted <TM>

F:269-291/Domain: Transmembrane #status predicted <TM>

F:305-328/Domain: Transmembrane #status predicted <TM>

F:123-203/Disulfide bonds: #status predicted

F:342/Binding site: palmitate (Cys) (covalent) #status predicted

F:372/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 31.0%; Score 469.5; DB 2; Length 381;

Matches 113; Conservative 75; Mismatches 123; Indels 53; Gaps 13.









QY 153 MKWQYTNRRRAFTMLGVVWVAVVIGSPMMHVOOL-----EIKYDFLYEKEHICGLEWT 206  
 Db 141 KQWPNKNNHACFGITVIGFAMACSTPLMMYSVLTDPEPKNISLDSYICK--YVCLEDDP 198  
 QY 207 SPVHOKIYTTFTILVILFLPLMVMILLYSKIGYEIWKRRVGGSGVLRTHGKEMSKAR 266  
 Db 199 EDKRLSYTTTLTILYLGPLGFIFVYCYTKI--FLRKKR-----NNMDKIRL 245  
 QY 267 KK ---KRAVIMMVTVALFAVCWAPFVHVMHMEYSNFEKEDDVTIKMFAIVOLI 320  
 Db 246 NKYSSTTKRINIMILISVVGFCALCPPEFENLVPDNH--EAVATCNHNLILFLICHLT 303  
 QY 321 GFNSNCGNPVYAFMNMENKKNVLSAVCY 350  
 Db 304 AMISTCVNPIFYGHIAKKNFORDILOFFENEC 333

## RESULT 12

S50150  
 gastric CCK-A receptor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 20-Apr-2000  
 C:Accession: S50150

R:Reuben, M.; Rising, L.; Prinz, C.; Hersey, S.; Sachs, G.  
 Biochim. Biophys. Acta 1219, 421-427, 1994  
 A:Title: Cloning and expression of the rabbit gastric CCK-A receptor.  
 A:Reference number: S50150; MUID:95002144; PMID:7918628

A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-427 <REP>  
 C:Superfamily: neurokinin 1 receptor

Query Match 18.0%; Score 402.5; DH 2; Length 427;  
 Best local Similarity 27.0%; Pred. No. 5, 6e-25;  
 Matches 95; Conservative 72; Mismatches 126; Indels 59; Gaps 7;

QY 49 VLTGVLIFALFGNALVYVYVTRSKAMRTVNIIFICLSLSDLLITFCIPVTMLQN 108  
 Db 45 ILVSLIFLISVIGNTLIVTVIRKMRMTVNIIFLSIALSDMLCIPMPFNIL 104  
 QY 109 DNMLGAFICKMVPFVQSTAVVTEILTMTCTIAVERHOGVHPHFKK-KQYTNRRRAFTMLG 167  
 Db 105 KDFIFGSAICAKTTLTYLMTTSVSVSTLNVLVAISLERYGAIKPKLQSPVWQ-TKSHALKVIA 163  
 QY 168 VVMVAVVIGSPMMHVOOLFIKYDFLYEKEHICGLEWTSPVHOKIYTTFTILVILFLPL 227  
 Db 164 ATWCLSPAIMTP-YPIYSNLVPTKTNTNNTANMKRFLPLSPVMMQQAHTLELLILFLIP 222  
 QY 228 MVMLILYSKIGYELW-----IKRRV----- 247  
 Db 223 LVMMVAYGMISLELYGCIKFDASOKKSAKERRKASTGCRPFENDCCYLQRSKPTTHOLFLQ 282  
 QY 248 ---GKSVLRTTHGKEMSKIAKKKKRAVIMMVTVALFVWVAFVHVMHMEYSNFEK 303  
 Db 283 QLSGGGGGRVSRTHSSSSAAALMAKKRVIRMLVWLVFLCWMPE-----IFSANAAR 335  
 QY 304 EYDDVITK-----MIFATVGLIGESNSICNPVYAFMNMENKKNVLSAVCY 350  
 Db 336 AYDVSAERLSTPISFTILLSYSSCVNPIIYCFMKNRRLGPMATFPCC 387

## RESULT 13

S17783  
 tachykinin receptor homolog DTKR - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C:Date: 22-Nov-1996 #sequence\_revision 22-Nov-1996 #text\_change 20-Apr-2000  
 C:Accession: S17783

R:Li, X.J.; Wolfgang, W.; Wu, Y.N.; North, R.A.; Forte, M.  
 EMBO J. 10, 3221-3229, 1991  
 A:Title: Cloning, heterologous expression and developmental regulation of a Drosophila  
 A:Reference number: S17783; MUID:92007772; PMID:1717263  
 A:Accession: S17784

A:Molecule type: mRNA  
 A:Residues: 1-519 <LIX>  
 A:Cross-references: FMB: X62711; NID: g8505; PIDN: CAA44595.1; PID: g8506  
 A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in lacking 481  
 C:Genetics:

A:Gene: FlyBase: Takr99D  
 A:Cross-references: FlyBase: FBgn0004622  
 C:Superfamily: neurokinin 1 receptor  
 C:Keywords: G protein-coupled receptor, neurotransmitter receptor, transmembrane prot

Query Match 17.9%; Score 401; DH 2; Length 519;  
 Best local Similarity 25.8%; Pred. No. 9, 2e-25;  
 Matches 87; Conservative 76; Mismatches 128; Indels 46; Gaps 11;

QY 47 ALVLTGVLIFALFGNALVYVYVTRSKAMRTVNIIFICLSLSDLLITFCIPVTMLQN 106  
 Db 105 SILFGCMVIVATG--CNLIVVIVMTTKMRIVTNYFVNIISIAADAMVSSINVTFNYYM 162  
 QY 107 ISDNWLGGAFCIKMVPFVQSTAVVTEILTMTCTIAVERHOGVHPHFKKQYTNRRRAFTML 166  
 Db 163 LDSLWMPFGFYCKLSQFIAMLSICASVFTLMAISIDRYVATIRPLQPR--MSKRCNLIAIA 220  
 QY 167 GVVMVAVVIGSPMMHVOOLFIKYDFLYEKEH-----CCLEFW-TSVH-----OK 212  
 Db 221 AVTLASTLISCPMM-----IYRIEVPVWGLSNVTYYPEWPGSPINHSIMES 279  
 QY 213 IYTFITLIVLFLPLMVMILLYSKIGYELWIKRVGGSGVLRTHGKEMSKIAKKKKAV 272  
 Db 271 LYNLILITUTYPLPIVSMVTYTSRVGFIAMGSKTICGCT-----PROENV-KSKRHVV 323  
 QY 273 IMMVTVALFVWVAFVHVMHMEYSNFEKEDDVT-----IKMIFATVLIQIGESN 328  
 Db 324 KMLVVLVIFALCWLDPH-----SYFIIITSCYPAITTEAPFIQRIYLAIVYLAAMSNN 377  
 QY 329 PIVVA-MNENKKNVLSAVCYCI-VNKTFSIPAQRHCN 364  
 Db 378 PIICWNNSRFRYGFKNVFRWCLFVRVGTSPSRREN 414

## RESULT 14

R40470  
 glucocorticoid-induced receptor precursor, short form Rp23 - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 20-Apr-2000  
 C:Accession: B40470

R:Harrigan, M.T.; Campbell, N.F.; Bourgeois, S.  
 Mol. Endocrinol. 5, 1331-1338, 1991  
 A:Title: Identification of a gene induced by glucocorticoids in murine T-cells: a pot  
 A:Reference number: A40470; MUID:92123228; PMID:1663214  
 A:Accession: B40470

A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-423 <HAR>  
 A:Cross-references: GB: M80481; GB: M80610; NID: q193516; PIDN: AAA17882.1; PID: q460318  
 C:Superfamily: neurokinin 1 receptor

Query Match 17.9%; Score 400.5; DH 2; Length 423;  
 Best local Similarity 27.4%; Pred. No. 8e-25;  
 Matches 85; Conservative 77; Mismatches 127; Indels 21; Gaps 6;

QY 45 KLALVLTGVLIFALFGNALVYVYVTRSKAMRTVNIIFICLSLSDLLITFCIPVTML 104  
 Db 71 KALLIVAYSFTIVSILFGNLVCHVIFKNORMHSATSLFVNLAVADIMITLNTPTPLV 130  
 QY 105 QNISDNWLGGAFCIKMVPFVQSTAVVTEILTMTCTIAVERHOGVHPHFKKQYTNRRRAFT 164  
 Db 131 REVNSTWVFGKGVCHVSFAQYCSLHVSALTILTAJAVDRHOVIMHPLKPRISIT--KGV 188  
 QY 165 MGVVMVAVVIGSPMMHVOOLFIKYDFLYEKEH-----CCLEFWTSP--VHOKIYTTFTILV 220  
 Db 189 YIAVIMVMTFPLSLPHAIICOKL---FTPKYSHDIVRSICLPDPFEPADLFWKYLDIAFTI 245  
 QY 221 ILFLPLVMYMLILYSKIGYELWIKKPVGTGSVLRTTHGKEMSKIAKKKKRAVIMMVTVA 280

```

Db      246 LLYLLPLFIISVARVAKKLWLCNTIGD-----VTTEOYLALRRKKKTIVKMLVLVVV 299
QY      281 LFAYCVAPFHVHVMHIIYSNFEKEDDVTIKMIFAIVQILGFSNSICNPVYAFPMNENPK 340
Db      300 LFALCWFLNCYVLLSSKAH-----TNNALYFAFHWFAMSSCTCNPFIYCWLNEFER 353
QY      341 KNVLSAVCYC 350
Db      354 VELKALLSMC 363

```

## RESULT 15

T29741

hypothetical protein C39E6.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Apr-2000

C:Accession: T29741

R:Fulton, B.; Le, T.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of C. elegans cosmid C39E6.

A:Reference number: 220676

A:Accession: T29741

A&gt;Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-457 &lt;FULL&gt;

A:Cross-references: EMBL J049944, FIDN AAA93419.1, CESP.C39E6.6

C:Genetics:

A:Gene: CESP.C39E6.6

A:Introns: 58/3; 107/1; 169/3; 253/3; 274/2; 329/2; 366/2; 438/2

C:Superfamily: neurokinin 1 receptor

```

Query Match      17.8%   Score 398;   DB 2;   Length 457;
Best Local Similarity 26.9%   Prod No 1.4e-24;
Matches 104;   Conservative 84;   Mismatches 131;   Indels 68;   Gaps 14;

```

```

QY      41 PGRAKLALV--LTGVL-IFALALFGNALVFYVYVTRSKAMRTVTNIFICSLALSDLITFF 97
Db      19 PSQSIYAIVPPLVIVLPLPFGVLTIVTCSHKAIIIVQNIPIILNLAASDCMCMCII 78
QY      98 CIPVTMLQNSDNWLGCAFICKKVPVFOSTAVVTEILTWTCTAVERHQGLVHPFKMKQY 157
Db      79 SLPIITINVKNNYFQNLCHLPCIQGISIEVCTFSLGAIADRYILVWRPHST--PL 136
QY      158 TNRAFTMLGVVWLVAVVIGSP-MWRVVOOLEIKYDELYEKEHIC---CLEEWTSPVHOKI 213
Db      137 SORCAFLLTVLLWLLILSPVVTLPYAFNMQIE-----YTEERICGYFCTEKWESAKSRA 190
QY      214 YTFILVILFLLPLMWMLIYSKI-----GYELWIKKRVGDGSLVLT-----IHG---- 258
Db      191 YTMIVMLAQFVVPFAVMAFCYANIVSVLSKRAQTKIKKVVERTSALFESSCAFPSHGLPQY 250
QY      259 -----KEMSKIAKKKRAVIMVTVVVALFVCAWAPHVHVHMMIYYSNFEKEY- 305
Db      251 ENELNEELDKQEKQKQVLLQNRRTTSILVTMVWEGITWLPHPHIVLSLIEYDDTQSFFP 310
QY      306 ----DDVTIKMIFAI-VQILGFSNSICNPVYAFPMNENPKKNVLSAVCYCIVNKTFSPAQ 360
Db      311 LYGRDDYDISYLLNLFTHSIFMSNNVLPVLYAWLPSPRQ-----LVIKIY---- 357
QY      361 RHNSGITMMRKKAQKSLRENPEETK 387
Db      358 -FGD-----RRKSDRIINQTSVYKTK 377

```

Search completed: June 10, 2003, 13:18:24  
Job time : 19 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:14:42 ; Search time 15 seconds  
(without alignments)  
1191 753 Million cell updates/sec

Title: US-10-070-241b-1  
Perfect score: 2235  
Sequence: 1 MQALNITPEQSRLLRDHNL..... RHLALFRSELAENSLDGH 431

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query % | Match | Length | DB         | ID     | Description  |
|------------|-------|---------|-------|--------|------------|--------|--------------|
| 1          | 508   | 22.7    | 417   | 1      | NFF2_RAT   | Q9eqd2 | rattus norv  |
| 2          | 502   | 22.5    | 522   | 1      | NFF2_HUMAN | Q9y5x5 | homo sapien  |
| 3          | 469.5 | 21.0    | 381   | 1      | NY2R_HUMAN | Q9q146 | homo sapien  |
| 4          | 468.5 | 21.0    | 381   | 1      | NY2R_MACMU | Q9qk74 | macaca mula  |
| 5          | 467.5 | 20.9    | 382   | 1      | NY2R_PIG   | O02836 | sus scrofa   |
| 6          | 467   | 20.9    | 384   | 1      | NY2R_BOVIN | P79113 | bos taurus   |
| 7          | 466.5 | 20.9    | 385   | 1      | NY2R_MOUSE | P97295 | mus musculus |
| 8          | 459.5 | 20.6    | 460   | 1      | OX2R_RAT   | P56719 | rattus norv  |
| 9          | 457   | 20.4    | 381   | 1      | NY2R_CAVPO | Q922d5 | cavia porce  |
| 10         | 455   | 20.4    | 444   | 1      | OX2R_HUMAN | Q43614 | homo sapien  |
| 11         | 451   | 20.2    | 444   | 1      | OX2R_CANFA | O9tup7 | canis fami   |
| 12         | 446.5 | 20.0    | 375   | 1      | NY4R_MOUSE | Q61041 | mus musculus |
| 13         | 446   | 20.0    | 371   | 1      | NY6R_RABIT | P79217 | oryctolagus  |
| 14         | 440   | 19.7    | 385   | 1      | NY2R_CHICK | Q9ddn6 | gallus gall  |
| 15         | 437   | 19.6    | 432   | 1      | NFF1_RAT   | Q9ep86 | rattus norv  |
| 16         | 432.5 | 19.4    | 430   | 1      | NFF1_HUMAN | Q9a266 | homo sapien  |
| 17         | 431.5 | 19.3    | 346   | 1      | GALR_RAT   | Q26805 | rattus norv  |
| 18         | 431   | 19.3    | 383   | 1      | NY1R_CAVPO | Q9wvd0 | cavia porce  |
| 19         | 430.5 | 19.3    | 348   | 1      | GALR_MOUSE | P56479 | mus musculus |
| 20         | 430.5 | 19.3    | 375   | 1      | NY4R_RAT   | Q63447 | rattus norv  |
| 21         | 429   | 19.2    | 371   | 1      | NY6P_MOUSE | Q61212 | mus musculus |
| 22         | 426   | 19.1    | 382   | 1      | NY1R_MOUSE | Q04573 | mus musculus |
| 23         | 425.5 | 19.0    | 375   | 1      | NY4R_HUMAN | P50391 | homo sapien  |
| 24         | 425.5 | 19.0    | 449   | 1      | NYR_DROME  | P25931 | rattus norv  |
| 25         | 425   | 19.0    | 382   | 1      | NY1R_RAT   | P21555 | rattus norv  |
| 26         | 424   | 19.0    | 383   | 1      | NY1R_PIG   | O02835 | sus scrofa   |
| 27         | 423.5 | 18.9    | 382   | 1      | NY1R_CANFA | Q02813 | canis fami   |
| 28         | 422.5 | 18.9    | 425   | 1      | OX1R_HUMAN | Q43613 | homo sapien  |
| 29         | 421.5 | 18.9    | 384   | 1      | NY1R_HUMAN | P25929 | homo sapien  |
| 30         | 420.5 | 18.8    | 349   | 1      | GALR_HUMAN | P47211 | homo sapien  |
| 31         | 412.5 | 18.5    | 416   | 1      | NY1R_XENLA | P34992 | xenopus lac  |
| 32         | 412.5 | 18.5    | 416   | 1      | OX1R_RAT   | P56718 | rattus norv  |
| 33         | 407   | 18.2    | 422   | 1      | GP72_CANFA | Q9ttq9 | canis fami   |

|    |       |      |     |   |            |        |              |
|----|-------|------|-----|---|------------|--------|--------------|
| 34 | 404   | 18.1 | 453 | 1 | CKKR_XENLA | P70031 | xenopus lae  |
| 35 | 401.5 | 18.0 | 423 | 1 | GF72_HUMAN | Q9aym4 | homo sapien  |
| 36 | 401   | 17.9 | 519 | 1 | TLR2_DROME | P30975 | drosophila   |
| 37 | 400.5 | 17.9 | 423 | 1 | GP72_MOUSE | P30731 | mus musculus |
| 38 | 393.5 | 17.6 | 427 | 1 | CKKR_RABIT | O97772 | oryctolagus  |
| 39 | 386.5 | 17.3 | 407 | 1 | NK1R_CAVPO | P30547 | cavia porce  |
| 40 | 384   | 17.2 | 402 | 1 | NK2R_CAVPO | Q64077 | cavia porce  |
| 41 | 383.5 | 17.2 | 407 | 1 | NK1R_HUMAN | P25103 | homo sapien  |
| 42 | 382.5 | 17.1 | 444 | 1 | CKKR_RAT   | P30551 | rattus norv  |
| 43 | 379   | 17.0 | 428 | 1 | CKKR_HUMAN | P32238 | homo sapien  |
| 44 | 378.5 | 16.9 | 430 | 1 | CKKR_CAVPO | Q69331 | cavia porce  |
| 45 | 378   | 16.9 | 385 | 1 | NK3R_MOUSE | P47937 | mus musculus |

## ALIGNMENTS

|          |  |           |      |         |
|----------|--|-----------|------|---------|
| RESULT 1 |  |           |      |         |
| ID       | NFF2_RAT   | STANDARD; | PRT; | 417 AA. |
| AC       | Q9EQD2;  |           |      |         |
| DT       | 16-OCT-2001 (Rel. 40, Created)   |           |      |         |
| DI       | 16-OCT-2001 (Rel. 40, Last sequence update)  |           |      |         |
| DE       | 16-OCT-2001 (Rel. 40, Last annotation update)  |           |      |         |
| DE       | Neuropeptide FF receptor 2 (Neuropeptide G protein-coupled receptor).  |           |      |         |
| GN       | NPGR OR NPFF2.   |           |      |         |
| OS       | Rattus norvegicus (Rat).   |           |      |         |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |           |      |         |
| OC       | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  |           |      |         |
| OX       | NCBI_TaxID=10116;  |           |      |         |
| RN       | [1]  |           |      |         |
| RP       | SEQUENCE FROM N.A.   |           |      |         |
| RC       | STRAIN=Sprague-Dawley;   |           |      |         |
| RX       | MEDLINE=20564301; PubMed=11024015;   |           |      |         |
| RA       | Reini J A , Jones K A , Adham N , Forray C , Artymyshyn R ,  |           |      |         |
| RA       | Durkin M M , Smith K E , Tamm J A , Boleju L W , Lakhani P P ,   |           |      |         |
| RA       | Raddatz R , Yao W - J , Goczalek K L , Boyle N , Kouranova E V ,   |           |      |         |
| RA       | Quan Y , Vaysse P J , Wetzel J M , Branchek T A , Gerald C ,   |           |      |         |
| RA       | Borowsky R ;   |           |      |         |
| RT       | Identification and characterization of two G protein-coupled   |           |      |         |
| RT       | receptors for neuropeptide FF.   |           |      |         |
| PL       | J Biol Chem. 275:39324-39331(2000).  |           |      |         |
| CC       | FUNCTION: RECEPTOR FOR NPAF (A-18-F-AMIDE) AND NPFF (F-8-F-AMIDE)  |           |      |         |
| CC       | NEUROPEPTIDES, ALSO KNOWN AS MORPHINE-MODULATING PEPTIDES. CAN   |           |      |         |
| CC       | ALSO BE ACTIVATED BY A VARIETY OF NATURALLY OCCURRING OR SYNTHETIC   |           |      |         |
| CC       | FMRF-AMIDE LIKE LIGANDS. THIS RECEPTOR MEDIATES ITS ACTION BY  |           |      |         |
| CC       | ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-  |           |      |         |
| CC       | CALCIUM SECOND MESSENGER SYSTEM.   |           |      |         |
| CC       | SUBCELLULAR LOCATION: Integral membrane protein  |           |      |         |
| CC       | SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  |           |      |         |
| CC       | HIGHEST TO OREXIN, NPY AND CHOLECYSTOKININ RECEPTORS.  |           |      |         |
| CC       | -----  |           |      |         |
| CC       | This SWISS-PROT entry is copyright. It is produced through a collaboration   |           |      |         |
| CC       | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |           |      |         |
| CC       | the European Bioinformatics Institute. There are no restrictions on its  |           |      |         |
| CC       | use by non-profit institutions as long as its content is in no way   |           |      |         |
| CC       | modified and this statement is not removed. Usage by and for commercial  |           |      |         |
| CC       | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |           |      |         |
| CC       | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  |           |      |         |
| CC       | -----  |           |      |         |
| DR       | EMBL AF268960, AAG41399.1, -   |           |      |         |
| DR       | InterPro: IPR000276; GPCR_Phodpsn.   |           |      |         |
| DR       | Pfam: PF00001; 7tm1; 1.  |           |      |         |
| DR       | PRINTS: PK00237; GPCRHHODPSN.  |           |      |         |
| DR       | PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.   |           |      |         |
| DR       | PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1  |           |      |         |
| KW       | G-Protein coupled receptor; Transmembrane; Glycoprotein.   |           |      |         |
| FT       | DOMAIN 1 45  |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 46 66   |           |      |         |
| FT       | DOMAIN 67 82   |           |      |         |
| FT       | CYTOPLASMIC (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 83 103  |           |      |         |
| FT       | DOMAIN 104 119   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 120 140   |           |      |         |
| FT       | DOMAIN 141 159   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 160 180   |           |      |         |
| FT       | DOMAIN 181 200   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 201 220   |           |      |         |
| FT       | DOMAIN 221 240   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 241 260   |           |      |         |
| FT       | DOMAIN 261 280   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 281 300   |           |      |         |
| FT       | DOMAIN 301 320   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 321 340   |           |      |         |
| FT       | DOMAIN 341 360   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 361 380   |           |      |         |
| FT       | DOMAIN 381 400   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 401 420   |           |      |         |
| FT       | DOMAIN 421 440   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 441 460   |           |      |         |
| FT       | DOMAIN 461 480   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 481 500   |           |      |         |
| FT       | DOMAIN 501 520   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 521 540   |           |      |         |
| FT       | DOMAIN 541 560   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 561 580   |           |      |         |
| FT       | DOMAIN 581 600   |           |      |         |
| FT       | EXTRACELLULAR (POTENTIAL)  |           |      |         |
| FT       | TRANSMEM 601 620   |           |      |         |

FT DOMAIN 141 160 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 161 181 4 (POTENTIAL).  
 FT DOMAIN 182 217 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 218 248 5 (POTENTIAL).  
 FT DOMAIN 249 274 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 275 295 6 (POTENTIAL).  
 FT DOMAIN 296 310 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 311 331 7 (POTENTIAL).  
 FT DOMAIN 332 417 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 118 206 BY SIMILARITY.  
 FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 417 AA: 47710 MW: DFE66AC63AF2AAD6 CRC64:  
 Query Match 22.7% Score 508; DB 1; Length 417;  
 Host local Similarity 31.8%; Pred. No. 5.9e-25;  
 Matches 119; Conservative 72; Mismatches 145; Indels 38; Gaps 9;  
 QY 54 LIPALALFQNALVYVYTRSKAMRTVTNIFIGSLALSLLITFFCIPVIMLQNSIDNWL 113  
 DB 54 LIPFLCMGNTVCFVWIRNRYMHTVTNFFINLAISDLLVGLFCMPITLLDNIAGWPF 113  
 QY 114 GAPICKMVPFVUSTAVVTEILMTCTIAVERHOGCLVHPKMKWQYTNRAFTMLGVWLVA 173  
 DB 114 GSSMCKISLGVGISVAASVFTLVAIVDRFCVVYPPKPK-LTVKTAFYMIIVIGLA 171  
 QY 174 VIVGSP MMHVOQLIKYDFEYKEH-----ICGLEEWTSPIVQIKYITFTILVFL 224  
 DB 172 ITIMTSAIMLHVQ--EKYVNVYRILSSHNKTSTVYVWCHEDWPNQHMRYITTVLPATYIL 229  
 QY 225 LPLMWMILYSLKIGELWTKKVGDSVLRITHGKMSKIARKKRAVIMVTVVVALFV 284  
 DB 230 APLSLVIMYARIGASLEKTSIAHSTGK--GRLEQWIVSKKKVQKIMULTYVALLFIL 284  
 QY 285 CMAPFHVVMHMLYENFKEKYDDVTIKMIFATVQIIGFSNSICNPVYVAFMNEFKKNVL 344  
 DB 285 SWLPPLWTLMKLSYADLSPNKLRLVINYVYFAHMLAFCSNVNPIIYGFENFRSGFQ 344  
 QY 345 SAVCYIVNKTSPAGRUHNSGITHMKKKAKESLR-----ENPVEETKG-RAFS 392  
 DB 345 DAFOFC-OKKKVPOFAYG-----LRAKRNIDINTSGLLVHPASQNPNGENIGCKKSA 396  
 QY 393 DGNIEVKLCEOTEE 406  
 DB 397 DNPTQESLMEETGE 410  
 RESULT 2  
 ID NFF2\_HUMAN STANDARD; PRT: 522 AA.  
 AC QYX55; Q9NR49;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Neuropeptide FF receptor 2 (Neuropeptide G protein-coupled receptor)  
 DE (G-protein-coupled receptor ILIAR77).  
 GN NPGPR OR NPPF2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (LONG ISOFORM). AND TISSUE-SPECIFICITY  
 RC MEDLINE-99180505; PubMed-10079187;  
 RA Cikos S., Gredor P., Koppel J.;  
 RT "Sequence and tissue distribution of a novel G-protein-coupled  
 RT receptor expressed prominently in human placenta";  
 RL Biochem. Biophys. Res. Commun. 256:452-456(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (SHORT ISOFORM), AND CHARACTERIZATION.

RC TISSUE-Brain;  
 RX MEDLINE-20408933; PubMed-10851242;  
 RA Elshourbagy N.A., Ames P.S., Fitzgerald L.B., Foley J.J.,  
 RA Chambers J.K., Szekeres P.G., Evans N.A., Schmidt D.B., Buckley P.T.,  
 RA Dyko G.M., Murodock P.R., Milligan G., Gloorke D.A., Tan K.B.,  
 RA Shabon U., Nuthalaganti P., Wang D.Y., Wilson S., Hergsma D.J.,  
 RA Sarau H.M.;  
 RT "Receptor for the pain modulatory neuropeptides FF and AF is an orphan  
 RT G protein-coupled receptor";  
 RL J. Biol. Chem. 275:25965-25971(2000).  
 [3]  
 RP SEQUENCE FROM N.A. (SHORT ISOFORM), AND CHARACTERIZATION.  
 RC TISSUE-Spinal cord;  
 RX MEDLINE-20564301; PubMed-11024015;  
 RA Bonini J.A., Jones K.A., Adham N.A., Forray C., Artymyshyn R.,  
 RA Durkin M.M., Smith K.E., Tamm J.A., Boteju L.W., Lakhani P.P.,  
 RA Raddatz R., Yao W.-J., Wojczalek K.L., Boyle N., Kouranova E.V.,  
 RA Quan Y., Vaysse P.J., Wetzel J.M., Branchek T.A., Gerald C.,  
 RA Borowsky B.;  
 RT "Identification and characterization of two G protein-coupled  
 RT receptors for neuropeptide FF";  
 RL J. Biol. Chem. 275:30324-30331(2000)  
 CC -1- FUNCTION: RECEPTOR FOR NPAF (A-18-F-AMIDE) AND NPEF (F-8-F-AMIDE)  
 CC NEUROPEPTIDES, ALSO KNOWN AS MORPHINE-MODULATING PEPTIDES. CAN  
 CC ALSO BE ACTIVATED BY A VARIETY OF NATURALLY OCCURRING OR SYNTHETIC  
 CC FMRE-AMIDE LIKE LIGANDS. THIS RECEPTOR MEDIATES ITS ACTION BY  
 CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-  
 CC CALCIUM SECOND MESSENGER SYSTEM.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: THE LONG ISOFORM IS ABUNDANT IN PLACENTA.  
 CC RELATIVELY HIGHLY EXPRESSED IN THYMUS, TESTIS, AND SMALL  
 CC INTESTINE. EXPRESSED AT LOW LEVELS IN SEVERAL TISSUES INCLUDING  
 CC SPLEEN, PROSTATE, BRAIN, HEART, OVARY, COLON, KIDNEY, LUNG, LIVER  
 CC AND PANCREAS AND NOT EXPRESSED IN SKELETAL MUSCLE AND LEUKOCYTES  
 CC HIGHEST BUT RELATIVELY LOW LEVEL OF THE SHORT ISOFORM IN PLACENTA  
 CC AND VERY LOW LEVEL IN NUMEROUS TISSUES INCLUDING ADIPOSE TISSUE  
 CC AND MANY BRAIN REGIONS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC HIGHEST TO OXPHIN, NPY AND CHOLEYSTOKININ RECEPTORS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF119815; RAD22047.1;  
 CC EMBL: AF257210; AAF67878.1;  
 CC EMBL: AF268899; AAG41398.1;  
 CC Genbank: HGNC:4525; GPR74.  
 CC InterPro: IPR000276; GPCR\_Rhodpsn.  
 CC Pfam: PF00001; 7tm\_1; 1  
 CC PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 CC PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 CC G-protein coupled receptor; Transmembrane; glycoprotein;  
 KW Alternative splicing.  
 FT DOMAIN 1 147 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 148 168 1 (POTENTIAL).  
 FT DOMAIN 169 184 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 185 205 2 (POTENTIAL).  
 FT DOMAIN 206 221 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 222 242 3 (POTENTIAL).  
 FT DOMAIN 243 262 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 263 283 4 (POTENTIAL).  
 FT DOMAIN 284 319 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 320 340 5 (POTENTIAL).  
 FT DOMAIN 341 377 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 378 398 6 (POTENTIAL).  
 FT DOMAIN 399 413 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 414 434 7 (POTENTIAL).  
 FT DOMAIN 435 522 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 220 308 BY SIMILARITY.  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1 102 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 466 466 A -> T (IN REF. 1). CRC64;  
 SQ SEQUENCE 522 AA; 60270 MW; 40CB9FCD42F77041 CRC64;  
 Query Match 22.5%; Score 502; DB 1; Length 522;  
 Best Local Similarity 30.1%; Pred. No 1 7e-24;  
 Matches 128; Conservative 87; Mismatches 162; Indels 48; Gaps 12;

QY 14 LLRDHNLREOFIALYRLPLVYTPPELPGRAKLAALVLTGLVLFALALFGNALFYVYVTRS 73  
 Db 128 LYSIDNIT---YVNYLHQPV-----AAFIISYELIFFLCMMGNTVYVCFVYMRN 175  
 QY 74 KAMRTVTNIFICSLALSLLITFRFPVIMLQNSDNMLGAFICKMPVFWQSTAVVTEI 133  
 Db 176 KHMVTNLFILNLAISDLVGFICMPITLLDNIIACMPFGNTMCKISGILVQGISVAASV 235  
 QY 134 LMTCTIAVERHGLVHPKMKQVYNNRRAFTMLGVWLVAVVIGSP---MHWVQVLEIKY 190  
 Db 236 FTLVAIAVDRECVVYPEPKR--LTIKTAFVIMIIWVLAITIMSPSAVMLHVQE-EKYY 292  
 QY 191 DFLEYKEH-----ICCLEWTSVPHOKIYTFILVILFLPLMWMLILYSKIGVYLWIKK 245  
 Db 293 RYRLNSQNTSPVYWCREDWPNQEMRKIYTVLFANILYLAFLSLIVIMYGRIGISLF--- 349  
 QY 246 RVGDSGLVRLTHGKEMSK-----IARKKRAVIMVTVVAVFVAVCAFPFHVHMMIYSNF 301  
 Db 350 ---RAAVPHT--GRKNQPOHVVSKKQKIKLLIIVALLFLLSWLPWLWMLSDVADL 404  
 QY 302 EKEYDDVTKIMFAIVQIIGFSNSICNPDIYAFNENFKNVLNLSAVCIYVNTFTSPAQR 361  
 Db 405 SPNELQIINIYIYFAHWAFLAGNSVNPDIYGFENFRGQFQAFQQLCQKRAKPMEA 464  
 QY 362 HNSGITIMWKKAKSLRENVEETKGEA-FSDGNIEVKLEQEPKPKIKRLHAIIPSE 420  
 Db 465 YA-----LKAKSHVLINTSNQLVQESTQNPFGHETLLYKSAEPQOE---LYWEE 512  
 QY 421 LAENS 425  
 Db 513 LKETT 517

## RESULT 3

NY2R\_HUMAN  
 ID NY2R\_HUMAN STANDARD, PRT, 381 AA.  
 AC P49146; Q13281; Q13457; Q9UE67;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neuropeptide Y receptor type 2 (NPY2-P) (NPY-Y2 receptor).  
 GN NPY2R.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi;  
 CC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=96070760; PubMed=7592910,  
 RA Gerald C., Walker M.W., Vaysse P.J., He C., Branche T.A.,  
 RA Weinschenk R.L.;  
 RT "Expression cloning and pharmacological characterization of a human  
 RT hippocampal neuropeptide Y receptor subtype.";  
 RL J. Biol. Chem. 270:26758-26761(1995).  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;

RX MEDLINE=96226058; PubMed=8632753;  
 RA Gehlert D.R., Beavers L.S., Johnson D., Gackenheimer S.L.,  
 RA Schober G.A., Galski P.A.;  
 RT "Expression cloning of a human brain neuropeptide Y Y2 receptor.";  
 RL Mol. Pharmacol. 49:224-228(1996).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96032678; PubMed=7559383;  
 RA Rose P.M., Fernandes P., Lynch J.S., Frazier S.T., Fisher S.M.,  
 RA Kodukula K., Kienzie B., Seethala R.;  
 RT "Cloning and functional expression of a cDNA encoding a human type 2  
 RT neuropeptide Y receptor.";  
 PL J. Biol. Chem. 270:22661-22664(1995).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96209788; PubMed=8643460;  
 RA Yan H., Yang J., Marasco J., Yamaguchi K., Brenner S., Collins F.,  
 RA Karbon W.;  
 RT "Cloning and functional expression of cDNAs encoding human and rat  
 RT pancreatic polypeptide receptors.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4461-4465(1996).  
 RN (5)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97131518; PubMed=8975716;  
 RA Aumar D.A., Badie D.M., Wong D.J., Ma Y.-Y., Kolakowski L.F. Jr.,  
 RA Yang-Feng T.L., Thompson D.A.;  
 RT "Characterization of the human type 2 neuropeptide Y receptor gene  
 RT (NPY2R) and localization to the chromosome 4q region containing the  
 RT type 1 neuropeptide Y receptor gene.";  
 RL Genomics 38:392-398(1996).  
 RN (6)  
 RP SEQUENCE FROM N.A.  
 RA Zastawny R.;  
 RT "Human neuropeptide Y Y2 receptor gene.";  
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE PANK  
 CC ORDER OF AFFINITY OF THIS RECEPTOR FOR PANCREATIC POLYPEPTIDES IS  
 CC PYY > NPY > PYY (3-36) > NPY (2-36) > [ILE-31, GLN-34] PP >  
 CC [LEU-31, PRO-34] NPY > PP, [PRO-34] PYY AND NPY FREE ACID.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN AMYGDALA, CORPUS CALLOSUM,  
 CC HIPPOCAMPUS AND SUBTHALAMIC NUCLEUS. ALSO DETECTABLE IN CAUDATE  
 CC NUCLEUS, HYPOTHALAMUS AND SUBSTANTIA NIGRA  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC HIGHEST TO TACHYKININS RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U36269; AAC50281.1; -;  
 DR EMBL: U42766; AAB04120.1; -;  
 DR EMBL: U32500; AAA93170.1; -;  
 DR EMBL: U42389; AAB07760.1; -;  
 DR EMBL: U50146; AAC51115.1; -;  
 DR EMBL: U76254; AAD00248.1; -;  
 DR Genbank: HGNC:7957; NPY2R.  
 DR MIM: 162642; -;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm1; 1.  
 DR KINIS: PK00247; GPCR\_HGLOPSN.  
 DR PROSITE: PS00337; G-PROTEIN\_RECEP\_FL\_1; 1.  
 DR PROSITE: PS00362; G-PROTEIN\_RECEP\_FL\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Phosphorylation, Lipoprotein; Palmitate.  
 FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 50 72 1 (POTENTIAL).  
 FT DOMAIN 73 82 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 83 104 2 (POTENTIAL).







```

Db      265  -ORRHKMTKMLVGVVVFVAVSWLPLHAFOLAVDIDSHVLDLKEY-----KLITVPHIA 318
QY      322  FSNSTCNPIVYAFMNNENFKKWLNAVYGVNKTFTSPAQR...HGNSGTTMPKRAKFS 377
Db      319  MCSTFANLLYCWNNNSNYKAFISA-----FRCEQRLDAIHSEVSMTEKAKK----- 365
QY      378  LRENPEETKGAESDGNIEVK 399
Db      366  -----NLEVK 370

RESULT 8
OX2R.RAT
ID OX2R.RAT STANDARD: PRT: 460 AA.
AC P5619;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).
GN HCRTR2.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N A
RC TISSUE=Brain;
RX MEDLINE=98150861; PubMed=9491997;
RA Sakurai T., Anemiyi A., Ishii M., Matsuzaki I., Chemelli R.M.,
RA Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
RA Arch J.R.S., Buckingham R.F., Haynes A.C., Carr S.A., Annan R.S.,
RA McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
RA Yanagisawa M.;
RT *Orexins and orexin receptors: a family of hypothalamic neuropeptides
RT and G protein-coupled receptors that regulate feeding behavior.*;
RL Cell 92:573-585(1998).
RN [2]
RP REVIEW.
RX MEDLINE=21237974; PubMed=11340621;
RA Hungs M., Mignot E.;
RT *Hypocretin/orexin, sleep and narcolepsy.*;
RL Bioessays 23:397-408(2001).
RN [3]
RP REVIEW.
PX MEDLINE=21178476; PubMed=11283317;
RA Willie J.T., Chemelli R.M., Sinton G.M., Yanagisawa M.;
RT *To eat or to sleep? Orexin in the regulation of feeding and
RT wakefulness.*;
PI Annu. Rev. Neurosci. 24:429-458(2001).
CC -|- FUNCTION: NONSELECTIVE, HIGH-AFFINITY RECEPTOR FOR BOTH OREXIN-A
CC AND OREXIN-B NEUROPEPTIDES.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- TISSUE SPECIFICITY: Expressed in the brain in the cerebral
CC cortex, septal nuclei, hippocampus, medial thalamic nuclei including
CC and median raphe nuclei, and many hypothalamic nuclei including
CC the tuberomammillary nucleus, dorsomedial hypothalamus,
CC paraventricular hypothalamic nucleus, and ventral premammillary
CC nucleus. Not detected in the spleen, lung, liver, skeletal muscle,
CC kidney and testis. Orexin receptor mRNA expression has also been
CC reported in the adrenal gland, enteric nervous system, and
CC pancreas.
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement! (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF041246; AAC40042 1;

```

```

DR INTERPT: IPK000276; GPCR_Rhodpsn.
DR InterPro: IP0004060; Orexin_receptor2.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PF00237; GPCRPHODPSN.
DR PRINTS: PF01522; OREXIN2P
DR PROSITE: PS00237; G-PROTEIN RECP_F1_1; 1.
DR PROSITE: PS00262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; transmembrane; Glycoprotein.
FT DOMAIN 1 54
FT TRANSMEM 55 75
FT DOMAIN 76 88
FT TRANSMEM 89 110
FT DOMAIN 111 127
FT TRANSMEM 128 150
FT DOMAIN 151 172
FT TRANSMEM 173 193
FT DOMAIN 194 224
FT TRANSMEM 225 247
FT DOMAIN 248 304
FT TRANSMEM 305 327
FT DOMAIN 328 342
FT TRANSMEM 343 366
FT DOMAIN 367 460
FT CARBOHYD 14 14
FT CARBOHYD 22 22
FT CARBOHYD 202 202
SQ SEQUENCE 460 AA; 52489 MW; 3B4E3D82F8B85D5 CRC64;
Query Match 20.6%; Score 459.5; DB 1; Length 460;
Best local Similarity 26.0%; Pred. No. 6,6e-22;
Matches 118; Conservative 97; Mismatches 164; Indels 75; Gaps 11;
QY 4 LNITPEQSRLLRDLNLTREOFIALYRLPLVITPELPGRKALALVLTGLVLFALALFON 63
DB 21 LNETQEPF---LNTDYDDEEFLR-YLWREYLIH-----PKYEYWLITAGYIIVFVALIGN 72
QY 64 ALVYVVTTRSKAMRTVINIFICSIALSDLLITFPICPVTMLQNSDNLGGAFCIKMVFP 123
DB 73 VLVCVAVKNNHUMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGSLCKVIPY 132
QY 124 VQSTAVVTEILTMTCIAVERHGLVHPFKMKQWYTNRAFTMLGVVWLNAVIVGSPMWIV 183
DB 133 LQTVSVSVSVSLTISCTALDRMYATICHPIMPK---STAKPAPNSIVTVIWSCTIIPQAIIV 190
QY 184 QULEIKYDFLYEKEHI--CCLEEMTSPVHQIKYITFELVLELLPLMVLMLYLSKIGYEL 241
DB 191 MERSMPLGLANKTTLTVCUERWGEVYPKMWHILCEFLVYMAPLCLVLAQLFKKL 250
QY 242 WIKKRVCDGSVLR-----T:HCKEMSKIARKKKRAVIMMVTV 278
DB 251 W-RQIPATSSVVQPKWKQPPVSGPPSGQSKAPISAVAAETKQIPAPPKTARMLMV-V 309
QY 279 VALFAVCWAPFHVHMMIE-YSNPEKEYDOVTIKMIFAIQVLIIGFSNSICNPVIYAFMNE 337
DB 310 LLVFAICVLPISILNVLKRVFGMTHTEDRETIVYAMFTFSHWLVLYANSAANPIYNYFUSG 369
QY 338 NFKKNVLNAVYGVNKNKTFSPAQRHNSGITMKKKAKEKSLRENPNVEETKGAESDGNIE 397
DB 370 KPEEFKFAFSCL-----GVHPPGCHPLAGRR-- 397
QY 398 VKLQHTQTEKKKLKRLHIALFR--SELAENSLIUS 429
DB 398 -----TSTESRSLTQISNFDNVSKLSEHVALTS 427
RESULT 9
NY2R.CAVPO
ID NY2R.CAVPO STANDARD: PRT: 361 AA.
AC Q922D5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Neuropeptide Y receptor type 2 (NY2R) (NPY-Y2 receptor).

```

GN NPY2R.  
 OS Cavia porcellus (Guinea pig).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 CC NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99017476; PubMed-9802430;  
 RA Sharma P.S., Holmberg S.K., Eriksson H., Beck-Sickinger A.G.,  
 RA Grundemar L., Larhammar D.;  
 RT "Cloning and functional expression of the guinea pig neuropeptide Y  
 Y2 receptor";  
 KL Requil. Rept. 75:23-28(1998).  
 CC -1- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC HIGHEST TO TACHYKININS RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF072821; AAD31443.1;  
 DR InterPro: IPR000276; GPCR\_Phodpsn.  
 DR Pfam: PF00001; 7tm.1; 1.  
 DR PRINTS: PR00247; GPCRHHODPSN.  
 DR PROSITE: PS00247; G-PROTEIN-RECEP\_FL1\_1; 1.  
 DR PROSITE: PS00247; G-PROTEIN-RECEP\_FL2\_1; 1.  
 DR PROSITE: PS00247; G-PROTEIN-RECEP\_FL3\_1; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Phosphorylation; Lipoprotein; Palmitate.  
 FT DOMAIN 1 50  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 51 73 1 (POTENTIAL).  
 FT DOMAIN 74 83  
 FT CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 84 105 2 (POTENTIAL).  
 FT DOMAIN 106 125  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 126 147 3 (POTENTIAL).  
 FT DOMAIN 148 167  
 FT CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 168 188 4 (POTENTIAL).  
 FT DOMAIN 189 215  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 216 241 5 (POTENTIAL).  
 FT DOMAIN 242 269  
 FT CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 270 292 6 (POTENTIAL).  
 FT DOMAIN 293 305  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 306 329 7 (POTENTIAL).  
 FT DOMAIN 330 381  
 FT CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 11 11 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT DISULFID 124 204  
 FT LIPID 343 343 PALMITATE (POTENTIAL).  
 SQ SEQUENCE 381 AA: 42342 MW: 4806859443 ECAC: CRC64;  
 Query Match 20.4%, Score 457, DB 1, Length 381.  
 Best Local Similarity 30.3%, Pred. No. 7.9e-22;  
 Matches 108; Conservative 75; Mismatches 12; Indels 50; Gaps 12;  
 QY 38 PELPGRAKL-----ALVLGVLLPALALFGNALVYVYVTRSKAMRTVNTNIFCSIALSDL 92  
 DB 49 PELIDSTKLTVEVVLILAYLSILLGVVGNLSLVHVIKFSKMRVTNTFFIANLAVADL 98  
 QY 93 LITFTFTPTVMTQNSDMWAGCAFLCKMVPFVQSTAVVTEILTMFCIAVERHOGIVHPFK 152  
 DB 99 LVNTLCPLTLYTLMGCKKMGCVLCHLVPAQGLAVOVSVITVIALDRHRCIVYHLD 158  
 QY 153 MKQYTNRAFTMLGVWLVAVVGVSPMHVQV-----LEIKYDFLYEKEHICCLEW---T 206  
 DB 159 SK--ISKNSPLIFGLAWGISALLAPLAFREYSLIFLPDF-----ELVACTEKWPCEE 212  
 QY 207 SPVHOKIYTFILVILPLPLVMVILLYSKICYELW--IKKRVGVGVSLRTIHCKEMSKI 264  
 DB 213 KSIYGVTVSLSLILYVLPGLITISVQYVRI-----WSKLNIVSPGAANTHYH----- 261

QY 265 ARKKKRAVIMVTVVAFVAVCWAPFHVHMMIEYNEF-----KEYDDVTIKMFAIVQIG 321  
 DB 262 -QRRQKTTKMLVFWVVVFAVSWLPIHAPOLAVDINSQVLDLKEY-----KLITVVFHIIA 315  
 QY 322 FSNISCNPIVYAFPMNENFKKNVLSAV-C-----YCIVNKTFSAPQHRGNSG 366  
 DB 316 MCSTFANPLLYGMNSNYRKAFLSAFRCQQLDAIQSEVCVCTGAKTNVEVEKNIG 371  
 RESULT 10  
 OX2R\_HUMAN  
 ID OX2R\_HUMAN STANDARD: PRT: 444 AA.  
 AC O43614;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).  
 GN HCRT2.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-98150861; PubMed-9491897;  
 RA Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,  
 RA Tanaka H., Williams S.C., Richardson J.A., Haynes A.C., Carr S.A., Annan R.S.,  
 RA Arch J.R.S., Buckingham R.F., Haynes A.C., Carr S.A., Annan R.S.,  
 RA McNulty D.E., Liu W.-S., Terrett J.A., Elshourbany N.A., Bergsma D.J.,  
 RA Yanagisawa M.;  
 FT "Orexins and orexin receptors, a family of hypothalamic neuropeptides  
 and G-protein-coupled receptors that regulate feeding behavior.";  
 KL Cell 92:573-585(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21580342; PubMed-11723285;  
 RA Oladunni R.P., Pye D.B., Scammell T.E., Matheson J.K.,  
 RA Stefansson K., Gulcher J.R.;  
 FT "Polymorphisms in hypocretin/orexin pathway genes and narcolepsy.";  
 KL Neurology 57:1896-1899(2001).  
 RN [3]  
 RP REVIEW.  
 RX MEDLINE-21237974; PubMed-11340621.  
 RA Hungs M., Mignot E.;  
 FT "Hypocretin/Orexin, sleep and narcolepsy.";  
 KL Bioessays 23:397-408(2001).  
 RN [4]  
 RP REVIEW.  
 RX MEDLINE-21178476; PubMed-11283317;  
 RA Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;  
 FT "To eat or to sleep? Orexin in the regulation of feeding and  
 wakefulness.";  
 KL Annu. Rev. Neurosci. 24:429-458(2001).  
 CC -1- FUNCTION: NONSELECTIVE, HIGH-AFFINITY RECEPTOR FOR BOTH OREXIN-A  
 CC AND OREXIN-B NEUROPEPTIDES.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF041245; AAC39602.1; .  
 DR EMBL; AY062031; AAL47215.1; .  
 DR Genew; HCN2:4849; HCRT2.  
 DR MIM; 602393; .  
 DR InterPro: IPR000276; GPCR\_Phodpsn.  
 DR InterPro: IPR004060; Orexin\_receptor2.



```

DR pfam: PF00001; 7tm_1;
DR PRINTS: PR00237; GPCRHHODPSN.
DR PRINTS: PR01522; OREXIN2R.
DR PROSITE: PS00237; G-PROTEIN_RECEPT_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECEPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 54 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 55 75 1 (POTENTIAL).
FT DOMAIN 76 88 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 89 110 2 (POTENTIAL).
FT DOMAIN 111 127 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 128 150 3 (POTENTIAL).
FT DOMAIN 151 172 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 173 193 4 (POTENTIAL).
FT DOMAIN 194 224 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 225 247 5 (POTENTIAL).
FT DOMAIN 248 304 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 305 327 6 (POTENTIAL).
FT DOMAIN 328 342 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 343 366 7 (POTENTIAL).
FT DOMAIN 367 444 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 14 14 N-LINKED (GLCNAC:...) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC:...) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC:...) (POTENTIAL).
SQ SEQUENCE 444 AA; 50680 MW; CA0669F0D4224C65 CRC64;

Query Match 20.4%; Score 455; DB 1; Length 444;
Best Local Similarity 27.7%; Pred. No. 1.2e-21;
Matches 104; Conservative 91; Mismatches 142; Indels 38; Gaps 9;

QY 4 LNIPEQSRLLRHNLTRQFIALYRLPLVYTPPELGRKALALVLTGVLIPALALFCN 63
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21 LNTEQEPF---LNPTDYDDEFLR-YLWREYVH----PKEYVWVLIAGYIIVFVALIGN 72
QY 64 ALVEVYVTRKAMRTVNTFCISLALSDLLITFCIPVTMLQNTSDMLGSAFCIKMVPF 123
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
73 VLVCVAVWKNHMRITVNTVFNLSADVLVITICLPATLVVDITETFWFFGQSICKVPI 132
QY 124 VQSTAVVTEILMTICIAVERHQGLVHPFKMKQWYNRAFTMLGVVWLVAVVGSPPMHV 183
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 LQTVSVSVSVLTSLALDRWYAICHPLMEK--STAKRARNISVILIIVSCIIMIPQAV 190
QY 184 QOLEIKYDFLYEKEHI--CCELEWTSYVHOKIYTFILVILFLLPLMVMILISKIYEL 241
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
191 MECSTVFPLGANKTTLFTVCDERWGGEIYPKMHICFFLVYMAPLCMLVLAQLFRKL 250
QY 242 WIKKRVGDSVLR-----TIHGKSKIAKKKKRAVIMMVTY 278
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 WCRQIPGTSSVVRKWKPIQVSPRGPGOPTKSRMSAVAAEIKQIPARKKTARMLMV-V 309
QY 279 VALFVAVWAPFVHVHMITE-YSNFEKEYDDVTIKMIFAIVQIGFSNSICNPVIYAFMNE 337
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
310 LLVFAICYLPISILNLVKRVGMEAHTEDETVYAVWFTFHWLVYANSAANPIYFNLSG 369
QY 338 NFKKNVLSAV-CYCI 351
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
370 KPREEFKAAFSCCCL 384

RESULT 11
OX2R_CANFA
ID OX2R_CANFA STANDARD; PRT; 444 AA.
AC Q9TUF7:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE OX2R receptor type 2 (Ox2r) (Hypocretin receptor type 2).
GN HCRTR2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=99385793; PubMed=10458611;
RA Lin L., Farago J., Li R., Kadotani H., Rogers W., Lin X., Qiu X.,
RA de Jong P.J., Nishino S., Mignot E.;
RT "The sleep disorder canine narcolepsy is caused by a mutation in the
RT hypocretin receptor 2 gene.";
RL Cell 98:365-376(1999).
RN [2]
RN REVIEW.
RX MEDLINE=21237974; PubMed=11340621;
RA Hungs M., Mignot E.;
RT "Hypocretin/orexin, sleep and narcolepsy.";
RL Histoassays 23:397-408(2001).
RN [3]
RN REVIEW.
RX MEDLINE=21178476; PubMed=11283317;
RA Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT "To eat or to sleep? Orexin in the regulation of feeding and
RT wakefulness.";
RL Annu. Rev. Neurosci. 24:429-458(2001).
RN [4]
RN VARIANT NARCOLEPSY LYS-54, AND MUTAGENESIS OF GLU-54.
RX MEDLINE=21189003; PubMed=11282968;
RA Hungs M., Fan J., Lin L., Lin X., Maki K.A., Mignot E.;
RT "Identification and functional analysis of mutations in the hypocretin
RT (orexin) genes of narcoleptic canines.";
RL Genome Res. 11:531-539(2001).
CC FUNCTION: N-SELECTIVE, HIGH-AFFINITY PEPTIDE-R P-P BOTH OREXIN-A
CC AND OREXIN-R NEUROPEPTIDES.
CC SURCELLULAR LOCATION: Integral membrane protein.
CC DISEASE: Defects in HCRTR2 are a cause of an autosomal recessive
CC form of narcolepsy, observed in Labrador, dobermans and
CC dachshunds. Narcolepsy is a neurological sleep disorder affecting
CC animals and humans, characterized by excessive daytime sleepiness,
CC sleep fragmentation, symptoms of abnormal rapid-eye-movement
CC (REM) sleep, such as cataplexy, hypnagogic hallucinations, and
CC sleep paralysis. Cataplexy is a sudden loss of muscle tone
CC triggered by emotions, which is the most valuable clinical feature
CC used to diagnose narcolepsy. As in humans, most cases of canine
CC narcolepsy are sporadic but an autosomal recessive form was also
CC observed.
CC SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF164626; AAD49333.1;
CC InterPro: IPR000276; GPCR_Rhodopsn
CC InterPro: IPR004060; Orexin_receptor2.
CC Pfam: PF00001; 7tm_1; 1.
CC PRINTS: PR00237; GPCRHHODPSN.
CC PRINTS: PR01522; OREXIN2R.
CC PROSITE: PS00237; G-PROTEIN_RECEPT_F1_1; 1.
CC PROSITE: PS50262; G-PROTEIN_RECEPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Disease mutation.
KW DOMAIN 1 54 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 55 75 1 (POTENTIAL).
FT DOMAIN 76 88 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 89 110 2 (POTENTIAL).
FT DOMAIN 111 127 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 128 150 3 (POTENTIAL).
FT DOMAIN 151 172 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 173 193 4 (POTENTIAL).
FT DOMAIN 194 224 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 225 247 5 (POTENTIAL).
FT DOMAIN 248 304 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 305 327 6 (POTENTIAL).

```

FT DOMAIN 128 442 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 343 466 7 (POTENTIAL).  
 FT DOMAIN 367 444 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 14 14 N-LINKED (GLCNAC: . .) (POTENTIAL).  
 FT CARBOHYD 22 22 N-LINKED (GLCNAC: . .) (POTENTIAL).  
 FT CARBOHYD 202 202 N-LINKED (GLCNAC: . .) (POTENTIAL).  
 FT VARIANT 54 54 E -> K (IN AUTOSOMAL RECESSIVE NARCOLEPSY).  
 FT MUTAGEN 54 54 E -> K: LOSS OF FUNCTION.  
 FT SEQUENCE 444 AA: 50675 MW: 108484.5360485066 CRC64:  
 Query Match 20.2%; Score 451; DB 1; Length 444;  
 Best Local Similarity 27.7%; Pred. No. 2,1e-21;  
 Matches 104; Conservative 90; Mismatches 143; Indels 38; Gaps 9;  
 QY 4 LNITPQFSRLRLDNLNITREQFIALYRLPLVYVTELPGRKALALVLTGLIFALALPQN 63  
 DB 21 LNITQEPF LNITDYDEDEFLR YLMREYLR ----PKYEWVLIAGYLIIVVVALVGN 72  
 QY 64 ALVYVYVTSKAMRTVNTIFGSLALSDLLITFFCIPVTMIQNSDNWIGAPFCKMVPF 123  
 DB 73 VLVCAVAKNHIMRTVNTVFNLSLADVLVITITCLPATLVVDITETWFGSLCKVTPY 132  
 QY 124 VQSTAVVTELLMTICIAVERHOGIVRPFKMKWQYTNRRAPTMLGVVWLVAVVIGSPMMHV 183  
 DB 133 IQTVSVSVSLHLSALDRDWTACHPLMFK--STAKRRNSIVITIVVSCIIIMIPQAV 190  
 QY 184 QLEIKYDFLYEKEHI--GCLLEWTSVPVHKIYTTFFILVFLFLPLVMYMLILYSKIYEL 241  
 DB 191 MECSLMPLGLANKTITFTWCDKRGGEIYPMHYICFFLVYMAPICLMVLAIVLIQIPKL 250  
 QY 242 WIKKRVGDSVLR -----TIHGKEMSKIAKKKRAVIMVTV 278  
 DB 251 WCRQIPGTSVVRKWKPLQPASQPRGQUTKRSISAAAEIKQIRARRKARMLMV-V 309  
 QY 279 VALFVACWAPFVHVHMMIE YSNFEKEYDDVTIKMIFAIVQIIGFSNLCNPVIVATPME 337  
 DB 310 LLVFAICYLPISILVAVKRVFQMPHTEDRETYYAWFTFSHWLVYANSAANPIIYNLSG 369  
 QY 338 NFKKNVLSAV-CYCI 351  
 DB 370 KPEEPKAAAFSCGCL 384

## RESULT 12

## NY4R\_MOUSE

ID NY4R\_MOUSE STANDARD: PRT: 375 AA.  
 AC Q61041;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Neuropeptide Y receptor type 4 (NPY4-R) (Pancreatic polypeptide receptor 1) (PPL) (NPYR D).  
 DE receptor 1) (PPL) (NPYR D).  
 GN PYPRI OR NPY4R.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV;  
 RX MEDLINE=9619393; PubMed=8641440;  
 RA Gregor P., Millham M.L., Feig Y., Decarr L.B., McCall M.L., Cornfield L.J.;  
 RT "Cloning and characterization of a novel receptor to pancreatic polypeptide, a member of the neuropeptide Y and peptide YY, the hank peptide, 381:58-62(1996)."  
 CC ORDER OF AFFINITY OF THIS RECEPTOR FOR PANCREATIC POLYPEPTIDES IS PP >> PYY >> NPY.  
 CC 1 SUBCELLULAR LOCATION: Integral membrane protein.  
 CC 1 TISSUE SPECIFICITY: HEART, DETECTED IN SMALL INTESTINE.  
 CC 1 SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS

CC CC  
 CC HIGHEST TO TACHYKININS RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC

DR EMBL: U40189; AAC52442.1; .  
 DR MGD: MGI:105374; Ppyrl.  
 DR InterPro: IPR003276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm.1; 1.  
 DR PRINTS: PR00237; GPCRHOPOPSN.  
 DR PROSITE: PS00247; G-PROTEIN\_RECEP\_FL\_1; 1.  
 DR PROSITE: PS00262; G-PROTEIN\_RECEP\_FL\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Phosphorylation; Lipoprotein; Palmitate.  
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 41 63 1 (POTENTIAL).  
 FT DOMAIN 64 74 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 75 95 2 (POTENTIAL).  
 FT DOMAIN 96 115 3 (POTENTIAL).  
 FT TRANSMEM 116 137 3 (POTENTIAL).  
 FT DOMAIN 138 157 4 (POTENTIAL).  
 FT TRANSMEM 158 178 5 (POTENTIAL).  
 FT DOMAIN 179 213 6 (POTENTIAL).  
 FT TRANSMEM 214 235 7 (POTENTIAL).  
 FT DOMAIN 236 265 8 (POTENTIAL).  
 FT TRANSMEM 266 288 9 (POTENTIAL).  
 FT DOMAIN 289 301 10 (POTENTIAL).  
 FT TRANSMEM 302 325 11 (POTENTIAL).  
 FT DOMAIN 326 375 12 (POTENTIAL).  
 FT CARBOHYD 2 2 N-LINKED (GLCNAC: . .) (POTENTIAL).  
 FT CARBOHYD 19 19 N-LINKED (GLCNAC: . .) (POTENTIAL).  
 FT CARBOHYD 29 29 N-LINKED (GLCNAC: . .) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC: . .) (POTENTIAL).  
 FT DISULFID 114 201 BY SIMILARITY.  
 FT LIPID 340 340 PALMITATE (POTENTIAL).  
 SQ SEQUENCE 375 AA: 42634 MW: 42634.17505310E CRC64:

Query Match 20.0%; Score 446.5; DB 1; Length 375;  
 Best Local Similarity 30.6%; Pred. No. 3.5e-21;  
 Matches 97; Conservative 72; Mismatches 121; Indels 27; Gaps 8.

QY 46 LALVLTGYLI-FALALFGNALVYVYVTSKAMRTVNTIFGSLALSDLLITFFCIPVTML 104  
 DB 41 LAFITTYSTETILGVNIGLICLIFFVTHQKKSNNVNLILANAFSDFLMCLICOPITVT 100  
 QY 105 QNISDNWLGGAFCIKWVPFVQSTAVTELLTMTCTINVERIUGLVHPFKMKWQYTNRRAPT 164  
 DB 101 YTIMDYIFGEVLCMKLTFIOCMSVTVSILSLVALERHQIINP--TGWKPSIFQAYL 158  
 QY 165 MLCGVVMVAVIVGSP-----MHWVQOQLEIKYDFLYEKEHICCLLEWTSVPVHOKIYT 215  
 DB 159 GIVVWIFSCFLSLPFLANSLNDFHYNHSKV-VELEDK--VVCVSVSSDHRLIYT 215  
 QY 216 TFIIVLFLPLVMYMLIYSKIGYELWIKKRVGDSVLRTHGKEMSKIAKKKRAVIMM 275  
 DB 216 TFIILFOYCIPLAFILVCYIRIYQRLQKHV-----PHAHACSSRAGQKRNMSL 267  
 QY 276 VIVVALFVACWAPFVHVHMMIEYSNFEKEYDDVTIKMIFAIVQIIGFSNLCNPVIVATPME 335  
 DB 268 MTMTAFVAVLPLHVFNTLEDW--YQEAIPACHGNLIFLMCHLLAMASTCVNPFYIGFL 325  
 QY 336 NENPKKNVLSAV--CYC 350  
 DB 326 NINPKKDIKALVLTGHC 342

## RESULT 13

## NY6R\_RABIT

ID NY6R\_PABIT STANDARD; PRT; 371 AA.  
AC P79217;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Neuropeptide Y receptor type 6 (NPY6-R).  
GN NPY6R.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Japanese white; TISSUE=Skeletal muscle;  
RX MEDLINE=9706688; PubMed=8910290;  
RA Matsumoto M., Nomura T., Momose K., Ikeda Y., Kondou Y., Akiho H.,  
RA Toqami J., Kimura Y., Okada M., Yamaguchi T.;  
RT "Inactivation of a novel neuropeptide Y/peptide YY receptor gene in  
RT primate species.";  
RL J. Biol. Chem. 271:27217-27220(1996).  
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE  
CC CYCLASE ACTIVITY.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC HIGHEST TO TACHYKININS RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D86521; BAA13104.1;  
DR InterPro: IPR000276; GPCR\_Rhodopsn.  
DR Pfam: PF00001; 7tm1; 1.  
DR PRINTS: P000237; GPCRHHODOPSN.  
DR PROSITE: PS00237; G-PROTEIN-RECEP\_FL1; 1.  
DR PROSITE: PS00262; G-PROTEIN-RECEP\_FL2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 34 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 35 57 1 (POTENTIAL).  
FT DOMAIN 58 69 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 70 90 2 (POTENTIAL).  
FT DOMAIN 91 110 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 111 132 3 (POTENTIAL).  
FT DOMAIN 133 152 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 153 174 4 (POTENTIAL).  
FT DOMAIN 175 208 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 209 230 5 (POTENTIAL).  
FT DOMAIN 231 261 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 262 284 6 (POTENTIAL).  
FT DOMAIN 285 297 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 298 321 7 (POTENTIAL).  
FT DOMAIN 322 371 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 109 196 HY SIMILARITY.  
FT LIPID 336 336 PALMITATE (POTENTIAL).  
SQ SEQUENCE 371 AA; 42447 MW; 5847E358EE320F24 CRC64;  
Query Match  
Best local similarity 27.6%; Score 446; DB 1; Length 371.  
Matches 100; Conservative 93; Mismatches 117; Indels 52; Gaps 13;  
QY 41 PGRKALVLGVLFALALFGRN ALFVYVYVTRSKAMRTVNTIFCSLASDILLTFCCI 99  
DB 31 PSIALLLLLIAYTVVLLMGICGNLSLITLIFKKQREAGQVNTNLIILANLSLSDIIVCVWCI 90

QY 100 PVTMLQNTSDNMLGAPICKMVPFVOSTAVVTEILMTCTIAVERHOCGLVHPFKMKWOYTN 159  
DB 91 PFTAIYTLMDRWIFGNTMCKLTYSVQSVISVSFSLVIAIERYGLIYNP--FGWKPSA 148  
QY 160 RRAETMLGVVWLVAIVIGSPM--WHV-----QQLEIKYDFLYEKEHICCLEWISPVHQ 211  
DB 149 SHAYWGIMLIWLFSLLSIPLLSYHLTDEPERNISLPTD-LY-SHRVVCVEHWPSKTNQ 206  
QY 212 KIYTFILVILFLPLVMILLYSKYELWIKKPV-HGSLVLTTHGCKEMSKIAFKK--- 268  
DB 207 LLYSTSLIMQYFVPLGFMEICYLKI-----VICLR-KENSKIDEPENE 250  
QY 269 -----KRAVIMVTVVAFVACWAPFHVHMMIEYSNFEKYDDVTIK---MIFAIWQ 318  
DB 251 SELTENKKINIMLSIVVTFACWLPNTENVIFDW-----YHEVLSMCHHDLVFAICH 304  
QY 319 IIGFSNCSINPIVYAMNENKKNVLSAVCYCIVNKTFFSPAQRHSGSI-TMMRKAKFS 377  
DB 305 LVAMVSTCINPLFYGLFNRRNFKDLVLIIHCL---CFALKRYENIAISTLIHDESKGS 361  
QY 378 LR 379  
DB 362 LR 363  
RESULT 14  
NY2R\_CHICK  
ID NY2R\_CHICK STANDARD; PRT; 385 AA.  
AC Q9DDN6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
GN Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RF SEQUENCE FROM N.A.  
RX MEDLINE=20532533; PubMed=11078884;  
RA Salanek E., Holmberg S K., Berglund M M., Roswell T., Latham D.,  
RT "Chicken neuropeptide Y receptor Y2: structural and pharmacological  
RT differences to mammalian Y2.";  
RL FEBS Lett. 484:229-234(2000).  
CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC HIGHEST TO TACHYKININS RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF309091; AAC37898.1;  
DR InterPro: IPR000276; GPCR\_Rhodopsn.  
DR Pfam: PF00001; 7tm1; 1.  
DR PRINTS: P000237; GPCRHHODOPSN.  
DR PROSITE: PS00237; G-PROTEIN-RECEP\_FL1; 1.  
DR PROSITE: PS00262; G-PROTEIN-RECEP\_FL2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 53 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 54 76 1 (POTENTIAL).  
FT DOMAIN 77 86 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 87 108 2 (POTENTIAL).  
FT DOMAIN 109 128 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 129 150 3 (POTENTIAL).



```

Db 69 LKNRHMRTVTNNFILNLAUSDLLVGFICMPTTLLVDNLITGWPEDNATCKMSGILVQMSVS 128
QY 131 TEILTMTCLAVRHOCILVHPFKMKWQYTNRRFTMLGVVWLAVIVGSPMHHVQOIEIKY 190
Db 129 ASVFTLVAIAVERFCIVHPFRK--ITLRKALFTIAIWAIAALLIMCPS-----AVTL 180
QY 191 DFLYEKEHI-----CCLEWTSPPHOKIYTFITILVILFLPLMVMMLILYSKIG 238
Db 181 TVTREHHFMLDARNRSYPIYSCHWEANPEKGMKVYIAVLFAHLYLVPLALIVVMYVKIA 240
QY 239 YELW-----IKRVGDGSLVIRTHGKEMSKTARKKKRAVIMMVTTVVALFVAVCWAPH 290
Db 241 RKLQOAGPARDTEEAAG-----GRTSRRRARVIMLVVALFTLSWPLW 289
QY 291 VVHMMIEYSNFEKEYDVTIKMI---FAIVQIIGFSNSICNPVIVAFMNEKKNVLSA 346
Db 290 VILLLLIDYG---ELSEIQLHLISVYAPPLAHWLAFFHSSANPIIYGYPNENPRGFOAA 345

```

Search completed: June 10, 2003, 13:17:17  
Job time : 16 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:15:17 Search time 35 Seconds  
(without alignments)  
2537.323 Million cell updates/sec

Title: US-10-070-241b 1

Perfect score: 2235

Sequence: 1 MQALNITPEQFSRLRDHNL.....RHLLAFRSELAENSPLDGSH 431

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvivirus.\*  
16: sp\_bacteriaph.\*  
17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Match | Length | DB ID     | Description         |
|------------|-------|-------|--------|-----------|---------------------|
| 1          | 1972  | 88.2  | 455    | 4 Q96P65  | Q96P65 homo sapien  |
| 2          | 496   | 22.2  | 408    | 4 Q96RV1  | Q96RV1 homo sapien  |
| 3          | 493.5 | 22.1  | 375    | 13 Q57463 | Q57463 brachydanio  |
| 4          | 485   | 21.7  | 417    | 11 Q924H0 | Q924H0 mus musculus |
| 5          | 484.5 | 21.7  | 374    | 13 Q97HX1 | Q97HX1 gadus morhua |
| 6          | 479.5 | 21.5  | 377    | 13 Q73733 | Q73733 brachydanio  |
| 7          | 479   | 21.4  | 381    | 11 Q96KQ0 | Q96KQ0 rattus norv  |
| 8          | 475.5 | 21.3  | 405    | 11 Q924N0 | Q924N0 mus musculus |
| 9          | 468.5 | 21.0  | 377    | 13 Q80GM3 | Q80GM3 gallus gall  |
| 10         | 452   | 20.2  | 373    | 13 Q73734 | Q73734 brachydanio  |
| 11         | 432.5 | 19.4  | 430    | 4 Q96RN3  | Q96RN3 homo sapien  |
| 12         | 431   | 19.3  | 385    | 13 Q80FM1 | Q80FM1 gallus gall  |
| 13         | 429   | 19.2  | 452    | 5 Q9VB87  | Q9VB87 drosophila   |
| 14         | 424   | 19.0  | 432    | 11 Q924G9 | Q924G9 rattus norv  |
| 15         | 422.5 | 18.9  | 425    | 4 Q9HVV6  | Q9HVV6 homo sapien  |
| 16         | 422   | 18.9  | 365    | 13 Q8UVW7 | Q8UVW7 lampetra il  |

## ALIGNMENTS

### RESULT 1

Q96P65 ID Q96P65 PRELIMINARY: PPT; 455 AA.  
AC Q96P65;  
DE 01-DEC-2001 (TRIMBLrel. 19, Created)  
DT 01-DEC-2001 (TRIMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TRIMBLrel. 21, Last annotation update)  
DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1).  
GN GPRI03.  
OS Homo sapiens (Human).  
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21458557; PubMed-11574155;  
RA Lee D K, Nguyen T, Lynch K W, Cheng P, Vanti W R, Arkhitko O,  
RA Lewis T, Evans J F, George S R, O'bowd H P;  
RT "Discovery and mapping of ten novel G protein-coupled receptor  
RT genes";  
RL Gene 275:83-91(2001).  
CC 1- FUNCTION: CYT-C OXIDASE IS THE COMPONENT OF THE RESPIRATORY  
CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNIT 1-  
CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE  
CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN  
CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2  
CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3  
CC AND COPPER B (BY SIMILARITY).  
CC 1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERROCYTOCHROME  
CC C + 2 H(2)O.  
CC 1- COFACTOR: HEMES A, A3, AND COPPER B (BY SIMILARITY).  
CC 1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN MITOCHONDRIAL  
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN MITOCHONDRIAL  
CC 1- INNER MEMBRANE (BY SIMILARITY).  
CC 1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
CC EMBL, AF411117, AAL26488.1; -;  
DR InterPro: IPR000883; COX1.  
DR TrEMBL: IPR002276; GPCR\_Phosphon.  
DR Pfam: PF00001, 7tm1, 1.

17 421.5 18.9 383 6 Q9GK75  
18 417 18.7 402 5 Q964E5  
19 410 18.3 372 11 Q9Z2D4  
20 408.5 18.3 375 6 Q97505  
21 408 18.3 429 5 P92045  
22 405.5 18.1 422 11 Q8VHD7  
23 404.5 18.1 431 5 Q8T8D1  
24 404 18.1 517 5 Q8TQ58  
25 400.5 17.9 504 5 Q9VCK8  
26 400.5 17.9 650 5 Q9VW75  
27 398 17.8 457 5 Q18534  
28 395 17.7 521 5 Q9VAD2  
29 394.5 17.7 481 5 Q967T7  
30 394 17.7 485 5 Q8SZ35  
31 390.5 17.5 678 5 Q94736  
32 390 17.4 465 5 Q44426  
33 387.5 17.3 540 5 Q9VRM0  
34 385 17.2 411 13 Q9W6I3  
35 384.5 17.2 370 13 Q8UWL5  
36 379 17.0 397 5 Q9NHA4  
37 377.5 16.9 423 5 Q964D4  
38 375.5 16.8 411 11 Q8R1P6  
39 375.5 16.8 412 5 Q20275  
40 375.5 16.8 436 11 Q8VCC7  
41 374 16.7 452 11 Q9JKN0  
42 373 16.7 463 11 Q9EPJ7  
43 370 16.6 398 4 Q9UDE6  
44 369.5 16.5 436 11 Q9DBV6  
45 368 16.5 380 13 Q9DGQ6

Q9GK75 macaca mula  
Q964E5 duessa tiq  
Q9Z2D4 cavia porce  
Q97505 sus scrofa  
P92045 lymnaea sta  
Q8VHD7 rattus norv  
Q8T8D1 urechis uni  
Q8TQ58 drosophila  
Q9VCK8 drosophila  
Q9VW75 drosophila  
Q18534 caenorhabdi  
Q9VAD2 drosophila  
Q967T7 drosophila  
Q8SZ35 drosophila  
Q94736 stomoxys ca  
Q44426 lymnaea sta  
Q9VRM0 drosophila  
Q9W6I3 gallus gall  
Q8UWL5 luqu rubrip  
Q9NHA4 boophilus m  
Q964D4 periplaneta  
Q8R1P6 cavia porce  
Q20275 caenorhabdi  
Q8VCC7 mus musculu  
Q9JKN0 mus musculu  
Q9EPJ7 mus musculu  
Q9UDE6 homo sapien  
Q9DBV6 mus musculu  
Q9DGQ6 carassius a

DR Pfam: PF00115; COX1; 1.  
 DR PRINTS: PR01570; NPFFRCEPTOR  
 DR PROSITE: PS00247; G\_PROTEIN\_RECEP\_FL\_1; UNKNOWN\_1.  
 DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_FL\_2; 1.  
 DR Copper; Heme; Inner membrane; Mitochondrion; Oxidoreductase; Receptor;  
 KW Respiratory chain; Transmembrane.  
 SQ SEQUENCE 455 AA; 51742 MW; A49H541CCFD36742 CRC64;

Query Match 88.2%; Score 1972; DB 4; Length 455;  
 Best Local Similarity 89.8%; Pred. No. 1e-157;  
 Matches 387; Conservative 0; Mismatches 6; Indels 38; Gaps 2;

QY 1 MQALNITPQPSKLLRDHNLTRQPTALYRLRPVYTPPLGKRAKIALVILCVLIFALAL 60  
 DB 63 MQALNITPEQSKLLRDHNLTRQPTALYRLRPVYTPPLGKRAKIALVILCVLIFALAL 122  
 QY 61 FGNALVEYVYVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLONISDNWLGGAFCICKM 120  
 DB 123 FGNALVEYVYVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLONISDNWLGGAFCICKM 182  
 QY 121 VPEVQSTAVVTELTMTCTIAVERHQGLVHPFKMKQYTNRRFTMLGVVWLVAIVVGSPPM 180  
 DB 183 VPEVQSTAVVTELTMTCTIAVERHQGLVHPFKMKQYTNRRFTMLGVVWLVAIVVGSPPM 242  
 QY 181 WHVQQLLEIKYDFLYEKHEICCELEWTSVPHQKIYTFILVILFLLPLMVMILLYSKICYE 240  
 DB 243 WHVQQLLEIKYDFLYEKHEICCELEWTSVPHQKIYTFILVILFLLPLMVMILLYSKICYE 288  
 QY 241 LWTIKKRVGDSVLTITIKKEMSK----IAPKKKRAVIMVTVVAFVACWAPFHVHMMIEYSN 300  
 DB 289 LWTIKKRVGDSVLTITIKKEMSK----IAPKKKRAVIMVTVVAFVACWAPFHVHMMIEYSN 324  
 QY 301 FEKEYDDVTIKMIFALVQIGFSNSICNPDIYVAFNENFKKNVLSAVCYIVNKTESPAQ 360  
 DB 325 FEKEYDDVTIKMIFALVQIGFSNSICNPDIYVAFNENFKKNVLSAVCYIVNKTESPAQ 384  
 QY 361 RHNSGITMMKKKAKFSLPENIVERTKGFATSDGNIYKLCGCTPEKKKLRHILALPRSE 420  
 DB 385 RHNSGITMMKKKAKFSLPENIVERTKGFATSDGNIYKLCGCTPEKKKLRHILALPRSE 444  
 QY 421 LAENSPDSGHI 431  
 DB 445 LAENSPDSGHI 455

RESULT 2  
 Q96RV1 PRELIMINARY; PRT; 408 AA.  
 AC Q96RV1;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE G-protein-coupled receptor 74.  
 GN GPR74.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Parker R.M., Herzog H.;  
 RT "Molecular cloning and characterization of GPR74 a novel G-protein  
 RT coupled receptor closest related to the Y-receptor family.";  
 RL Brain Res. Mol. Brain Res. 0:0-0(2000).  
 DR EMBL: AF236084; AAK58513.1;  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm1.1;  
 DR PRINTS: PR01570; NPFFRCEPTOR.  
 DR PROSITE: PR01572; NPFFRCEPTOR.  
 DR PROSITE: PS00247; G\_PROTEIN\_RECEP\_FL\_1; UNKNOWN\_1.  
 DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_FL\_2; 1  
 KW Receptor.  
 SQ SEQUENCE 408 AA; 47488 MW; EA1A70E4108C95; CRC64;

Query Match 22.2%; Score 496; DB 4; Length 408;  
 Best Local Similarity 30.2%; Pred. No. 1.2e-33;  
 Matches 123; Conservative 80; Mismatches 136; Indels 68; Gaps 13;

QY 14 LURDNLTRQPTALYRLRPVYTPPLGKRAKIALVILCVLIFALALFGNALVYVYVTRSR 73  
 DB 29 LXSNDINIT---VNVYLHQPV-----AAFTLSYPLIFFLCMGNTVVCVFVVRN 76  
 QY 74 KAMRTVTNIFICSLALSDLLITFFCIPVTMLONISDNWLGGAFCIKWVPFVQSTAVVTEI 133  
 DB 77 KHMHTVTNIFILNALSILVGLFCMPITLIDNLTACMPFGNTMCKISGLVQGSVAASV 136  
 QY 134 LWTCTIAVERHQGLVHPFKMKQYTNRRFTMLGVVWLVAIVVGSPP---MWHVQQLLEIKY 190  
 DB 137 FTLVALAVDFQCVVYPEPK--LTIKTAFVILMIIVLWLAITMSPSAVMLIHVE-EKYY 193  
 QY 191 DFLYKEH-----ICCLEWTSVPHQKIYTFILVILFLLPLMVMILLYSKICYEILWIKK 245  
 DB 194 RVLANSQNKTSVYVWCREIWPNOEMRKIYTVTVKANIYLAFLSLIVIMYGRIGISLP--- 250  
 QY 246 PVGDSVLTITIKKEMSK----IAPKKKRAVIMVTVVAFVACWAPFHVHMMIEYSNF 301  
 DB 251 ---RAAVPT---GRKNQFQWVSVKOKIIMKILIVALLIFLISMLPLTLMMLSDYADL 305  
 QY 302 FEKEYDDVTIKMIFALVQIGFSNSICNPDIYVAFNENFKKN----- 342  
 DB 306 SPNELQINIYIYFPAHLAFGNSSVNPDIYGFENENFRGFGEAFQQLCKPAKPMEA 365  
 QY 343 -VLISAVCYIVNKTESPAQ-----HCNSGITMMKKKAKFSLRNP 382  
 DB 366 YTLKAKSHVLTISNQLVQESTFQPHGCT--ILYKSA-----ENP 405

RESULT 3  
 O57463 PRELIMINARY; PRT; 375 AA.  
 ID O57463;  
 AC O57463;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Neuropeptide Y /peptide YY receptor Yb.  
 GN NPYRYH.  
 OS Brachydanio rerio (Zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lundell I., Berqlund M.M., Starback P., Salaneck S., Gehlert D.R.,  
 RA Larhammar D.;  
 RT "Cloning and characterization of a novel neuropeptide Y (NPY) receptor  
 RT subtype in the zebrafish.";  
 RL DNA Cell Biol. 0:0-0(1997).  
 DR EMBL: AF030245; AAB94616.1;  
 DR 2FIN: ZDB-GENE-980526-208; npyryb.  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR Pfam: PF00001; 7tm1.1;  
 DR PRINTS: PR00237; GPCR\_Rhodopsn.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_FL\_1; UNKNOWN\_1.  
 DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_FL\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 375 AA; 42116 MW; 4R0C9551131A14D1 CRC64;

Query Match 22.1%; Score 493.5; DB 13; Length 375;  
 Best Local Similarity 30.1%; Pred. No. 1.8e-33;  
 Matches 113; Conservative 79; Mismatches 142; Indels 41; Gaps 10;

QY 48 LVLTGLVIFALAFGNALVYVYVTRSKAMRTVTNIFICSLALSDLLITFFCIPVTMLONI 107  
 DB 28 LIAVYSTMLAVGLVGNLCVIVVITRKPMRNVNIFVNI:SCSDILVGLVCLPVTIYTL 87



QY 108 SDNMLGAFICKMVPVQSTAVVTEILTMTICIAVERHOGILVHPFKMKWOYTNNRAFTMLG 167  
 Db 88 MDRWILGEALCKVTPVQCMSTVTSFMSWLIALERHQLIHP--TGMVPVVRHSYLAVA 145  
 QY 168 VVWLIVAVIVSPMMHVOOL-----EIKYDFLYEKEHCLEEWTSVPVHOKIYTTFTILVI 221  
 Db 146 VIMLIACFISLPFLSFLNTSPFNLSLPFPNSDFHFCIEQWPSGNRLTYTTTLTLLAC 205  
 QY 222 LFLPLVWMLLLYSKIGYELWIKKRVGDSVLTITIGKEMSKIAKKKRAVIMVTVVAL 281  
 Db 206 QYCLPLALILVYFRIPLRLSRKDM-----VEPAPGGRQ--KKAQSKRVNAMLASIVAA 259  
 QY 282 FAVCAWPFHVHVMHMYSPNEK-----EYDDVTIKMIFAIVQIIGFNSICNPVAFMNE 337  
 Db 260 FALCWLPLNVENTIFDW-NHEALPVQCHD-----AIFSACHLTAMASTCVNPVYIGFLNN 313  
 QY 338 NFKKNVLSAVCYCIVNKTSPSPAQRHNSGITHMRKKAKFSLENPVVEFTKGEAFSDNGTE 397  
 Db 314 NPOKEKSLLSRSC---RCWGPARESY-----ESFPLSTVSTGITHKGSILNSGSAS 359  
 QY 398 VKLCEQTEKKKLLK 412  
 Db 360 T---YQPHKKNLSLEQ 371

## RESULT 4

Q924HO PRELIMINARY; PRT; 417 AA  
 AC Q924HO  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Neuropeptide NPFF receptor.  
 GN GPR74.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,  
 RA Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallico J.,  
 RA Wang R., Evans J., Gould P., Austin C.P.;  
 RT "Identification and characterization of two cognate receptors for  
 RT mammalian FMRFamide-like neuropeptides."  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF330054; AAK94198.1;  
 DR MGD: MG1:1860130; Gpr74.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR01570; NPFFRECEPTOR.  
 DR PROSITE: PS01572; NPFFRECEPTOR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 417 AA; 47485 MW; AC0686514CDIAF40C CRC64;

Query Match 21.7%; Score 485; DB 11; Length 417;  
 Best Local Similarity 31.0%; Pred. No. 1.le-32;  
 Matches 121; Conservative 71; Mismatches 136; Indels 62; Gaps 11;

QY 17 DHNLTROQFIALYRLPLVYTPPELPGRKALALVITGVLFALFALGNALVFFVVTTSKAM 76  
 Db 29 DINIT--YNYXVHQPQV-----AAVFSSYLLIFVLCMGVNTVVCIVIRNRHM 76  
 QY 77 RTVTNIFCSIALSDLLITPFCIPVTMLQNSIDNWLGGAFCICKMVPVQSTAVVTVITM 136  
 Db 77 HTVTNFEILNATSDLLVIGFCWPIITLLDNIAGWFGSGSMCKISGLVQGISVAASVFTL 136  
 QY 137 TCIAVERHOGILVHPFKMKWOYTNNRAFTMLGVVWLVAVIVGSP---MHWVQQLKIDYEL 193  
 Db 137 VAIAVDRFCVVPVFPKPK--LTVKTAFTVITVIGLAIATMTPSALMLHVQ--EEKYRV 192

QY 194 YEKEH-----ICCLEWTSVPVHOKIYTTFTILVILFLLPLVWMLLLYSKIGYELWIKRV 247  
 Db 193 RLSSHNTSTVYWCPCDWPHEMPPIYTTVLTATYTLAPLSLIVIMYAPIGASLF----- 247  
 QY 248 GDGSLVLTIRHG---EMSKIARKKKRAVIMVTVVVALFVCAWPFHVHVMHMYSPNEKE 304  
 Db 248 ---KTAHCTGKQRPQDWHVSKKKQKVIKMLLTVALFILSWLPWTLMMLSDYIDLSPN 304  
 QY 305 YDDVTIKMIFAIVQIIGFNSICNPVYAFMNEFKKNVLSAVCYCIVNKTSPSPAQRHNS 364  
 Db 305 KLRTIINYIYFPAHWAFCUNSSVNP-IYGFENFERNFGQDAFQIC----- 350  
 QY 365 SGITMPEKKAK---EFLSR--ENPVEETKG 388  
 Db 351 -----QKKAKPOEAYSRLAKRNKVVINTSG 374

## RESULT 5

Q9YHX1 PRELIMINARY; PRT; 374 AA.  
 AC Q9YHX1  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Neuropeptide Y-peptide YY receptor Yb.  
 GN NPYRB.  
 OS Gadus morhua (Atlantic cod).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99017376; Pubmed=9802392;  
 RA Arvidsson A.K., Wraith A., Jonsson-Rylander A.C., Larhammar D.;  
 RT "Cloning of a neuropeptide Y-peptide YY receptor from the Atlantic  
 RT cod: the Yb receptor."  
 RL Regul. Pept. 75:39-43(1998).  
 DR EMBL: AF073925; AAD02833.1;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GPCR\_Rhodopsn.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 374 AA; 42260 MW; D2A9C9516C4998E2 CRC64;

Query Match 21.7%; Score 484.5; DB 13; Length 374;  
 Best Local Similarity 27.2%; Pred. No. 1e-32;  
 Matches 108; Conservative 94; Mismatches 144; Indels 51; Gaps 10;

QY 23 EQFI-----ALYRLPLVYTPPELPGRKALIA---LVLGVLIFALFALGNALVFFVVTTS 73  
 Db 6 DOFIMESHPKANYSLIQLANDQECSPSSKGTFTLILVYSTWIAVGVNCSLVFVIARQ 65  
 QY 74 KAMRTVTNIFCSIALSDLLITPFCIPVTMLQNSIDNWLGGAFCICKMVPVQSTAVVTVI 133  
 Db 66 KEMHNTNIFIANISCDILMCIICLPVLTIVTLMRWILGEALCKLFPVQCISVTVSI 125  
 QY 134 LTMTCIAVERHOGILVHPFKMKWOYTNNRAFTMLGVVWLVAVIVGSPMH-----VQ 185  
 Db 126 FSLVLIAMERYQLIIHP--TGMKPMVQCQSYMAVGIILWVACILISVPFLSTVLDNLPLQN 183  
 QY 186 LEIKYDFLYEKEHCLEEWTSVPVHOKIYTTFTILVILFLLPLVWMLLLYSKIGYELWIKK 245  
 Db 184 LSLPFP---QODHWLCTESWPTNSRLAYTTSLLVFPYPLGLILAAACYLSIFLRLRRK 240  
 QY 246 RVCDGSLVLTIRHOKEMSKIAKKKRAVIMVTVVVALFVCAWPFHVHVMHMYSPNEKE 305  
 Db 241 -----DMVERARDSSRDNRKAGSRINVMGLSIVALFALFVCLPLNFTNFTEDW-----H 289  
 QY 306 DDVTI-----KWIFAIVQIIGFNSICNPVYAFMNEFKKNVLSAVCYCIVNKTSPSPAQR 361

```

Db 290 HELMVSCQINLIFSVC:HLVAMAST:VNPVVY:FLANSNFKQLKAILSH:---RCWGAAP 346
QY 362 HGNSGITMKKKAKESLRENPNVETKGEAFSDGNIEV 398
Db 347 YENLPIASV - STEVTKESHMSKGSISL 372

RESULT 6
074733 PRELIMINARY: PRT: 377 AA.
AC 074733:
DT 01-AUG-1998 (TEMBLrel. 07, Created)
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Neuropeptide Y/peptide YY receptor Ya.
GN NPYRYA OR NPYRYA.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE-98068842: PubMed-9407007;
RA Lundell I., Berglund M.M., Starback P., Salaneck E., Gehlert D.R.,
Larhammar D.;
RT Cloning and characterization of a novel neuropeptide Y receptor
subtype in the zebrafish *;
RL DNA Cell Biol. 16:1457-1463(1997).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE-98096493: PubMed-9434780;
RA Ringvall M., Berglund M.M., Larhammar D.;
RT Multiplicity of neuropeptide Y receptors: cloning of a third distinct
subtype in the zebrafish *;
RL Biochem. Biophys. Res. Commun. 241:749-755(1997).
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE-99337783;
RA Starback P., Lundell I., Fredriksson M., Berglund M.M., Yan Y.L.,
Wraith A., Soderberg C., Postlethwait J.H., Larhammar D.;
RT Neuropeptide Y receptor subtype with unique properties cloned in the
zebrafish: the zY receptor *;
RL Brain Res. Mol. Brain Res. 70:242-252(1999).
DR DMHL: AF037400; AAC41276 1;
DR HSSP: P02699; 1F88
DR ZFIN: ZDB-GENE-980526-393; npyrya.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm1; 1.
DR PRINTS: PR00237; GPCR_RHO_DOPSN.
DR PROSITE: PS00247; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 377 AA: 42901 MW: 60862AD9C7BFE5D0 CRC64;

Query Match 21.5%; Score 479.5; DB 13; Length 377;
Best Local Similarity 29.7%; Pred. No. 2.7e 32;
Matches 105; Conservative 83; Mismatches 131; Indels 35; Gaps 7;

QY 46 LALVLGVLFALALFAGNALPVVYVTRSKAMRTVTNIFCSIALSDLLITFFCIPVMTLQ 105
Db 43 LTLVLGYCLVLGLGLGILLICITIMQRPDPNVTSLILANLSVSDILVSFCLPFTVWY 102
QY 106 NTSNMGIGAFICKKVPQVSFAVVTEILITICIAVERHOGLVHPFKMKQYFNRRAPT 165
Db 103 TLMDDHIFALLCRMLPVCVSVTVSVLSVLIALERHQLHP---SGWKSPVQAYIA 160
QY 166 LGVWNLVAVIVGSPM: WH-----VQOLEIKYDFLYEKEHICCLEEWTSPVH 210
Db 161 VLTWLLACVTSPLAPHLITSGYSLFPAHLSQIQV-----CLEWVWPSQDH 208
QY 211 QKIYITFLVLFLPLMVLMLLYSKIGYEIWKPKVQSGSVLTPIHGKMSKIAKKR 270

```

```

Db 209 KLAYTTSLLFQVCGPCLLLMLLCLYLRIFLRRLRMLERQ:SR--NPEDHPRVMIHSKR 266
QY 271 AVIMKVTVALAVCHAPFHVHVMMI EYSNFKFKYDDV-TIKMIPAIVQIIGFSNSICNP 329
Db 267 INVMIAIATAAFVACWLPINAFNVV---ADCOEVLIPVCNHNHLLPSLCHILAMSTCVNP 323
QY 330 IYAFMNFNFKKNVLNSAVCYIVNKTFTSPAQRHKNSGITMMRKKAKEKSLRNPV 383
Db 324 ILYGFLSNFPRKDVASVVLHCHFPQPLEDSYRHPFPMSTMTDVSRISEFKLNNSV 377

RESULT 7
QYERCO PRELIMINARY: PRT: 381 AA.
ID QYERCO:
AC QYERCO:
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Neuropeptide Y/peptide YY-Y2 receptor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-SMALL INTESTINE;
RA Voisin T., Goumain M., Laburthe M.;
RT Cloning of a cDNA encoding a rat type 2 neuropeptide Y/peptide YY
receptor expressed in intestinal epithelial crypt cells *;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF004257; AAF89094.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR InterPro: IPR001220; Lectin_legh.
DR Pfam: PF00001; 7tm1; 1.
DR PRINTS: PR00237; GPCR_RHO_DOPSN.
DR PRINTS: PR01570; NPFFRECEPTOR.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 381 AA: 42510 MW: 659327904H288BC7 CRC64;

Query Match 21.4%; Score 479; DB 11; Length 381;
Best Local Similarity 29.9%; Pred. No. 3.1e 32;
Matches 120; Conservative 78; Mismatches 144; Indels 60; Gaps 12;

QY 17 DHNLTRQPIALYRLRPLVYTPPELPG-----RAKIALVITGVILFALALFGN 63
Db 9 DENQTVKVELYSGGPTTPRGELPPDPPELIDSTKLVEVGVVLLAYCSILLGVGN 68
QY 64 ALVYVYVTRSKAMRTVTNIFCSIALSDLLITFFCIPVMTIONISDNMGAFICKKVP 123
Db 69 SILVHVIVKFKSMRTVTNFIANLAVADLLVNTLCLDPTFTLYTLMGKWKGVILCHVPY 128
QY 124 VGSTAVTEILTMTCIAVERHOGLVHPFKMKQYFNRRAPTMTLGVVVLVAVIVGSPMHV 183
Db 129 AGCLAVGVSTITLVIALDRHRCIVYHLEK--ISKQISFLIIGLANCVSAILASPLAIF 186
QY 184 QO----LEIKYDFLYEKEHICCLEW---TSPVHOKIYTTFTLYILFLPLIPMVLMLLYSKI 237
Db 187 REYSLIEITPDP---EIVACTEKWPGEEKSVYGVTSLSLTLLIYVLPGLIISFSTRI 242
QY 238 GYEIM--IKKRVGDSGLVRLIHGKMSKIAKKKRAVIMVTVVALFACWAPFHVHMM 295
Db 243 -----MSKLNHVSPCAASDHYH-----QRRHKTKMLVCVVVVFVAFVSWLDIHAFOIA 290
QY 296 IEYSNFE---KEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFNFKKNVLSAVCYIV 352
Db 291 VDIIDSHVLDLKHV-----KLIFTVFHIIAMGSTEANPLLYCWNNSYKKAFLSA----- 339
QY 353 NKTFTSPAQR-----HCNSGITMMRKKAKEKSLRENPNVETKGEA 390
Db 340 ---FRCEQRDLAIHSEVSMTEFKAKKNLEVKKNGLTDSFSEA 378

```







```

QY 11 FSLRLKDHNLTRQFIATYLRRLRVVYTPELPGRAKALALVLTGVLIFALALFNAIFVYV 70
Db 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
32 FSSYCHSSPVAAMFIAV- - - - -VLFLFLCIVGNTLVFVIV 68
QY 71 TRSKAMKVTNIFGSLALSULLITFFCIPVTMLQNSIDNMLGSAFICKMVPFVGVSTAVV 130
Db 71 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 LKNRHMKVTNMFILMLAVSDI.PVGIFGMPFTPLVNLITGMPFONATCKMSGIVQGSVS 128
QY 131 TELLTTCIAVERHIGLVIHPEKMKQYTNRRAFITMLGVVMLVAVIVGSPMHVVOOLEIKY 190
Db 131 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 ASVETLVALAVERFCIVHPEK- - -LTLKALFTIAVIALALLIMCPS- - - - -AVTL 180
QY 191 DFLYEKEHI- - - - -OCLDEWTSVHQKIYTTFLVILFLPLMVLMLIISKIG 238
Db 191 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 TVTREHHPMLDARNRSPLYSCWGAMPEKMKRYTAVLFAHIVLPLALIVMYVRIA 240
QY 239 YELM- - - - -IKKRVGSLVRLTIHGKENSIAKRRKRAVIMVTVVVALFVAVCPFH 290
Db 239 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 RKLCCAGHVAROTEFARVAG- - - - -GRTSRRKRVVHMLAMVAFPTLSLPLW 289
QY 291 VVIMMIEYSNFEKEDDTIKI- - - - -FAIVQIIGFSNSICNPVYAFMNEKKNVLSA 346
Db 291 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
290 VLLALLIDYG- - - - -ELSELQHLLSVYAFPLAHHLAFAHSSANPIIYGYNENFRGFQA 345

RESULT 15
Q9H9V6 PRELIMINARY: PPT: 425 AA
AC Q9H9V6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE Hypocretin receptor-1 (Hypocretin receptor 1).
GN HCRTR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20429525; PubMed=10973418;
RA Peyron C., Faraco J., Rogers W., Ripley B., Overcom S., Charnay Y.,
RA Nevsimilova S., Aldrich M., Reynolds D., Albin R., Li R., Hungs M.,
RA Pedrazzoli M., Padigaru M., Kucherlapati M., Pan J., Maki R.,
RA Lammets G.J., Bouras C., Kucherlapati R., Nishino S., Mignot E.;
RT "A mutation in a case of early onset narcolepsy and a generalized
RT absence of hypocretin peptides in human narcoleptic brains.*";
RL Nat. Med. 6:991-997(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Yeager M., Welch R., Haque K., Bergen A.;
RX MEDLINE=21580342; PubMed=11723285;
RA Olatodotir B.R., Rye D.B., Scammell T.E., Matheson J.K.,
RA Stetansson K., Gulcher J.R.;
RT "Polymorphisms in hypocretin/orxin pathway genes and narcolepsy.*";
RL Neurology 57:1896-1899(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=202084; PubMed=11723285;
RA FMBL; AF202084; AAG28020.1; JOINED.
DR FMBL; AF202078; AAG28020.1; JOINED.
DR FMBL; AF202079; AAG28020.1; JOINED.
DR FMBL; AF202080; AAG28020.1; JOINED.
DR FMBL; AF202081; AAG28020.1; JOINED.
DR FMBL; AF202082; AAG28020.1; JOINED.
DR FMBL; AF202083; AAG28020.1; JOINED.
DR FMBL; AF070269; AAL50221.1; JOINED.
DR FMBL; AF062030; AAL47214.1; JOINED.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00247; GPCRHHODOPSIN.
DR PRINTS: PR01570; NPFRRECEPTOR.

```

```

DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 425 AA: 47535 MW: 8650B37FA2A096 PRC64;
Query Match 18.9%; Score 422.5; DH 4; Length 425;
Best local Similarity 26.8%; Pred. No. 1.9e-27;
Matches 106; Conservative 94; Mismatches 153; Indels 43; Gaps 10;
QY 23 EQFIALYRLRLPVLYTPELPGRAKALALVLTGVLIFALALFNAIFVYVYVTRSKAMRTVTNI 82
Db 23 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
29 DFLRL-YIMRDVLY- - - - -PKQYEWVLIAYAVVAVVAVLGNLTLCVLAAMKHHMRTVTNY 83
QY 83 FICSLALSULLITFFCIPVTMLQNSIDNMLGSAFICKMVPFVGVSTAVVTLTLMICIAVE 142
Db 83 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 FIVNLSADVLTALCLPASLLVDITESTWLFEGHALCKVPIYLOAVSVSAVLTLSFALD 143
QY 143 RHQGLVHPKMKQYTNRRAFITMLGVVMLVAVIVGSPMHVVOOLEIKYDFLYEKEHI- - -C 200
Db 143 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
144 RWATICHPIILPK- - -STARRARGSIILGIWASLIAMVPOAAVMEGSSVLPILANRTRLSV 201
QY 201 CLEEWTSVHQKIYTTFLVILFLPLMVLMLIISKIGYELWIK- - - - -KRV 247
Db 201 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
202 CLKRWADLLYPKLYHSCFFIVTYLAPLGLMAMAFYQIFPKLWGRQIIPGTISALVVRNKKRP 261
QY 248 GD- - -GSVLRTHCK- - - - -HMSKIARKKKRAVIMVTVVVALFVAVCPFHVHMM 295
Db 248 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
262 STPLD-LLEWGLSGEP- - - - -PKAKAFLAEVKQMPAPPPTAKMLW-VLLVLTALYLFISLVNL 329
QY 296 IR-YSNFKEKYDDVTIKMIFALVQIIGFSNSICNPVYAFMNEKKNVLSAVCYCIVN- 353
Db 296 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
321 KRVGMPHQAQDREAVYACFTSHMLVYANSAAP: IYNFISGKFRPQPKAATSCCIPGL 380
QY 354 - - - - -KTFSPAQRHSGITMRKKKAKFSLRNPV 383
Db 354 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
381 GPCGSLKAFSPHSSASHKSLSLQSCSISKISEHV 416

```

Search completed: June 10, 2003, 13:18:00  
Job time : 38 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:16:27 ; Search time 16 seconds  
(without alignments)  
792,580 Million cell updates/sec

Title: US-10-070-241b-1

Perfect score: 2235

Sequence: 1 MQALNITPEQFSRIIRHNI.....PHIALPRSELAENSPIDSGH 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID              | Description         |
|------------|-------|-------------|--------|--------------------|---------------------|
| 1          | 502   | 22.5        | 420    | US-09-255-368-6    | Sequence 6, Appli   |
| 2          | 479   | 21.4        | 381    | US-08-687-355A-4   | Sequence 4, Appli   |
| 3          | 479   | 21.4        | 381    | US-09-407-367-4    | Sequence 4, Appli   |
| 4          | 474   | 21.2        | 381    | US-08-687-355A-6   | Sequence 6, Appli   |
| 5          | 474   | 21.2        | 381    | US-09-407-367-6    | Sequence 6, Appli   |
| 6          | 469.5 | 21.0        | 381    | US-08-192-288-2    | Sequence 2, Appli   |
| 7          | 469.5 | 21.0        | 381    | US-08-687-355A-2   | Sequence 2, Appli   |
| 8          | 469.5 | 21.0        | 381    | US-09-200-673-16   | Sequence 16, Appli  |
| 9          | 469.5 | 21.0        | 381    | US-09-407-367-2    | Sequence 2, Appli   |
| 10         | 468.5 | 21.0        | 381    | US-08-876-798A-2   | Sequence 2, Appli   |
| 11         | 455   | 20.4        | 444    | US-09-426-290-2    | Sequence 2, Appli   |
| 12         | 453   | 20.3        | 444    | US-09-119-788-2    | Sequence 2, Appli   |
| 13         | 437   | 19.6        | 432    | US-09-255-368-2    | Sequence 2, Appli   |
| 14         | 432.5 | 19.4        | 430    | US-09-255-368-8    | Sequence 8, Appli   |
| 15         | 431.5 | 19.3        | 346    | US-09-199-737-5    | Sequence 5, Appli   |
| 16         | 431.5 | 19.3        | 346    | US-08-993-088A-3   | Sequence 3, Appli   |
| 17         | 431.5 | 19.3        | 346    | US-08-993-424B-3   | Sequence 3, Appli   |
| 18         | 431.5 | 19.3        | 346    | US-09-058-333A-5   | Sequence 5, Appli   |
| 19         | 430.5 | 19.3        | 348    | US-08-513-974B-46  | Sequence 46, Appli  |
| 20         | 430.5 | 19.3        | 348    | US-08-513-974B-342 | Sequence 342, Appli |
| 21         | 430.5 | 19.3        | 348    | US-08-993-088A-10  | Sequence 10, Appli  |
| 22         | 430.5 | 19.3        | 348    | US-08-993-424B-10  | Sequence 10, Appli  |
| 23         | 430.5 | 19.3        | 348    | US-08-540-650B-2   | Sequence 2, Appli   |
| 24         | 430.5 | 19.3        | 350    | US-08-495-648B-32  | Sequence 32, Appli  |
| 25         | 429   | 19.2        | 371    | US-08-415-818-6    | Sequence 6, Appli   |
| 26         | 429   | 19.2        | 371    | US-08-894-236-6    | Sequence 6, Appli   |
| 27         | 429   | 19.2        | 371    | US-08-919-624-4    | Sequence 4, Appli   |

## ALIGNMENTS

### RESULT 1

US-09-255-368-6

; Sequence 6, Application US/09255368

; Patent No. 6262246

; GENERAL INFORMATION:

; APPLICANT: Gerald, Christophe P.G.

; APPLICANT: Jones, Kenneth A.

; APPLICANT: Bonini, James A.

; APPLICANT: Borowsky, Beth

; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors

; FILE REFERENCE: 1795/57155-A

; CURRENT APPLICATION NUMBER: US/09/255,368

; CURRENT FILING DATE: 1999-02-22

; EARLIER APPLICATION NUMBER: 09/161,113

; EARLIER FILING DATE: 1998-09-25

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn Ver. 2.0 - beta

; SEQ ID NO 6

; LENGTH: 420

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-255-368-6

Query Match 22.5%, Score 502, Pos 4, Length 420;

Best Local Similarity 30.1%, Pred. No. 5, Pos 35;

Matches 128, Conservative 87, Mismatches 162; Indels 48; Gaps 12;

|    |     |  |     |
|----|-----|--|-----|
| QY | 14  | LRCHNLTPEQFTALYKRLPLVVTPELPKPAKALALVLIIVLIFALALEGNALVEYVYVTR   | 73  |
| DB | 26  | LYSDINIT---YNNYLRHQPV-----AAFIISYELFFICMGNVTVCIVMKN            | 73  |
| QY | 74  | KAMRIVNIFICSIALSDLLITFCITIVIMLQNIUNWAGAPICKMVPFVUSTAVVTEI      | 133 |
| DB | 74  | KIMHITVNLFLNLIAISDLVGIFCMFPTILLDNIIAGWPGNMTKISGLVQGISVAASV     | 133 |
| QY | 134 | LTMTCTIAVERHGGIVHPFKMKWQYTNKRATMLGVVWIVAVIVGSP---MHWVQQLIKY    | 190 |
| DB | 134 | FTVIAIVDRPQCVVYPFKPK--LITKTAFTVIMIIWVLAITIMSPSAVMIVQPF-EKY     | 190 |
| QY | 191 | DFLEYKEH-----ICCLEEWTSPVHQKIYTFITFLIFLFLMVMILLYSKIGYEILWIKK    | 245 |
| DB | 191 | RVKLNSQKTSPPVYWCREDPNQPMRKIVITVIFANILYIAPISLIVIMYGRIGISLIF---  | 247 |
| QY | 246 | RVGDSVLRTHIGEMSK-----IARKKKRAVIMWITVWALFVACWAPFVWMMHMYSNF      | 301 |
| DB | 248 | ---RAAVPHIT--GRKNQEQMIVYSRRKKQIKIKMLLVALLLFLLSWLFLWIMLMLSDYAL  | 302 |
| QY | 302 | EKEYDQVVIKMIFAIVQVIGFSNSICNPVIVYATMNIENKKNVLSAVCYCIIVAKIFSPAQR | 361 |

Db 103 SPNELGQINIIYIPFAHMLAFGNSSVNPILYGFNNFRRHQFQFAFQIQICOKRAKPMEA 362  
 QY 362 HGNSGITMMKKKAKFSLRNPVEETKGEA FSKNIEVKLCQETEEKKKLKPILLALPERSE 420  
 Db 363 YA-----LKAASHVLINTSNQLOVESTFONPIGETILLYRKSAEKPOQE-----LYMEE 410  
 QY 421 LAINS 425  
 Db 411 LRETT 415

## RESULT 2

US-08-687-355A 4  
 : Sequence 4, Application US/08687355A  
 : Patent No. 598944  
 : GENERAL INFORMATION:  
 : APPLICANT: Synaptic Pharmaceutical Corporation  
 : TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE  
 : TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF  
 : NUMBER OF SEQUENCES: 27  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cooper & Dunham LLP  
 : STREET: 1185 Avenue of the Americas  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: U.S.A.  
 : ZIP: 10036  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/687,355A  
 : FILING DATE: NO. 598944ember 26, 1996  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: White, John P.  
 : REGISTRATION NUMBER: 28,678  
 : REFERENCE/DOCKET NUMBER: 44742-A PCT/JPW/MAT  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 212-278-0400  
 : TELEFAX: 212-391-0525  
 : INFORMATION FOR SEQ ID NO: 4:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 381 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-687-355A 4

Query Match 21.4%; Score 479; DB 2; Length 381;  
 Best Local Similarity 29.9%; Pred. No. 4,le-33;  
 Matches 120; Conservative 78; Mismatches 144; Indels 60; Gaps 12;  
 QY 17 DHNITREOFIALYRLRPLVYTPPELPG-----RAKLAIVITGVLIIFALALFGN 63  
 Db 9 DENQTVKVELYSGSPTTPRGELPPEPELIDSTKLVEQVVLILAYCSILLGVGN 68  
 QY 64 ALVFFVYVTRSKAMRTVTNIFICSLASLIDLLITFCIPVTMLQNSDNLGGAFICKKVPF 123  
 Db 69 SLIVHVVVKFSKMRVITNFIANLAVADLVNLTLCIPFTITLYTMGMGKMGCPVICHLPY 128  
 QY 124 VOSTAVTEILMTCTIAYERHUGLVHPEKMKQVYNNRAFTMLGVVVLVAVVIGSPMHV 183  
 Db 129 AGLAVQVSTITLTVALDRHRCIVYHLESK--ISKQISFLITGLANGVSALLASPLAIF 186  
 QY 184 QO---LEIKYDPLYEKHEKICCLEEM---TSPVHOKIYTFITLITLPLLMVMIILYSKI 237  
 Db 187 REVSLEIITPDE--EIVACTEKWPGEKSVGYTVVYSLTLLILYVLPGLIISFSYTRI 242  
 QY 248 GYELW--TKKRVGCDGSLVLTTHGKEMSKIAKKKRAVIMMVTVVVALFVAVCWAFPHVHHM 295

Db 243 ----WSKLNKHNVSPOGASDHYH-----QRRHKHTTKMLVCVVVVFVAVSWLPLHAFOLA 290  
 QY 296 IEVSNFE---KEVDVETIKMIFAIQIITGFNSICNEIYVAFMNEKKNVLSAVYCIY 352  
 Db 291 VDIIDSHVLDKEY-----KLITVPHIILAMCTANPLLYGWMNSNTRKAFLSA----- 339  
 QY 353 NKTFSIPAOR-----HGNSGITMMKKKAKFSLRNPVEETKGEA 390  
 Db 340 ---FRCEQRDLAIHSEVSMTEFKAKKNLEVKKNNGSLTDSSEA 378

## RESULT 3

US-09-407-367-4  
 : Sequence 4, Application US/09407367  
 : Patent No. 6420532  
 : GENERAL INFORMATION:  
 : APPLICANT: Christophe P.G. Gerald, et al.  
 : TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2 SPECIFIC COM  
 : NUMBER OF SEQUENCES: 27  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cooper & Dunham LLP  
 : STREET: 1185 Avenue of the Americas  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: U.S.A.  
 : ZIP: 10036  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/407,367  
 : FILING DATE:  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: White, John P.  
 : REGISTRATION NUMBER: 28,678  
 : REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPW  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 212-278-0400  
 : TELEFAX: 212-391-0525  
 : INFORMATION FOR SEQ ID NO: 4:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 381 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-09-407-367-4

Query Match 21.4%; Score 479; DB 4; Length 381;  
 Best Local Similarity 29.9%; Pred. No. 4,le-33;  
 Matches 120; Conservative 78; Mismatches 144; Indels 60; Gaps 12;  
 QY 17 DHNITREOFIALYRLRPLVYTPPELPG-----RAKLAIVITGVLIIFALALFGN 63  
 Db 9 DENQTVKVELYSGSPTTPRGELPPEPELIDSTKLVEQVVLILAYCSILLGVGN 68  
 QY 64 ALVFFVYVTRSKAMRTVTNIFICSLASLIDLLITFCIPVTMLQNSDNLGGAFICKKVPF 123  
 Db 69 SLIVHVVVKFSKMRVITNFIANLAVADLVNLTLCIPFTITLYTMGMGKMGCPVICHLPY 128  
 QY 124 VOSTAVTEILMTCTIAYERHUGLVHPEKMKQVYNNRAFTMLGVVVLVAVVIGSPMHV 183  
 Db 129 AGLAVQVSTITLTVALDRHRCIVYHLESK--ISKQISFLITGLANGVSALLASPLAIF 186  
 QY 184 QO---LEIKYDPLYEKHEKICCLEEM---TSPVHOKIYTFITLITLPLLMVMIILYSKI 237  
 Db 187 REVSLEIITPDE--EIVACTEKWPGEKSVGYTVVYSLTLLILYVLPGLIISFSYTRI 242  
 QY 238 GYELW--TKKRVGCDGSLVLTTHGKEMSKIAKKKRAVIMMVTVVVALFVAVCWAFPHVHHM 295  
 Db 243 ----WSKLNKHNVSPOGASDHYH-----QRRHKHTTKMLVCVVVVFVAVSWLPLHAFOLA 290



QY 296 IYNSFE---KEYDDVTIKMIFALVQLIGFNSICNPVYAFMNFNPKKNVLSAVCYCIV 352  
 Db 291 VDISHVLDLKEY-----KLFTVFHIIAMCSTFANPLLYGWNNSYRKAFLSA----- 339  
 QY 353 NKTFSPAQR-----HGNSGITMMRKAKFSLRENPVEETKGEA 390  
 Db 340 ---FRCQKJLUAHSEVSMIFKAKKNLEVKKNCLILUSFSEA 378

## RESULT 4

US-08-687-355A-6  
 ; Sequence 6, Application US/08687355A  
 ; Patent No. 5989834  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Synaptic Pharmaceutical Corporation  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE  
 ; TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/687,355A  
 ; FILING DATE: No. 5989834ember 26, 1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 44742-A-PCT/IPW/MAT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-391-0525  
 ; TELEFAX: 212-391-0525  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 381 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-687-355A-6

Query Match 21.2%, Score 474; DR 2; Length 381;  
 Best Local Similarity 29.6%, Pred. No. 11e-32;  
 Matches 119; Conservative 78; Mismatches 145; Indels 60; Gaps 12;  
 QY 17 DHNLTREOFIALYRLRLVLTPELPG-----RAKIALVLTGVLIFALALPQN 63  
 Db 9 DENOTVEKVFYSGGTPTRGELPPDPELIDSTKLVEQVVLILAYCSILLGVGN 68  
 QY 64 ALVFVWVTRSKAMRTVNIIFCSLALSDDLITFCIPVTMLQNSDNLGGAFCIKMVPF 123  
 Db 69 SLVTHVVIKPKSMVIVNFFIANLAVADLLVNLICLPTLTITLTMGEKMGCVLCHLVY 128  
 QY 124 VOSTAVVTEILTMTCIAVERHQGLVHPFKMKQVYNNRAFTMLGVVWLVAIVGSPMHW 183  
 Db 129 AGGLAVQVSTITLVIALDRHRCIVYHLESK--ISKQISFLLIIGLAWGVSALLASPLAIF 186  
 QY 184 QO---LEIKYDFLYEKHEHCCLPEW---TSPVQKQIYTTTFLVILPLMWMILYLSKI 237  
 Db 187 REYSLIEIIPDF---EIVACTEKWGEKSVYGVTSLSLTLLVYLPGLGISFSYTRI 242  
 QY 238 GYELW--IKKRVGDSVLETHGKEMSKIAKKKKRAVINMMVTVALFVAGWAFPHVIMM 295  
 Db 243 ---WSKLNKHNVSFCAASDHYH-----ORRHKKTKMLVCVVVFAVSWLPLHAFOLA 290

QY 296 IYNSFE---KEYDDVTIKMIFALVQLIGFNSICNPVYAFMNFNPKKNVLSAVCYCIV 352  
 Db 291 VDISHVLDLKEY-----KLFTVFHIIAMCSTFANPLLYGWNNSYRKAFLSA----- 339  
 QY 353 NKTFSPAQR-----HGNSGITMMRKAKFSLRENPVEETKGEA 390  
 Db 340 ---FRCQKJLUAHSEVSMIFKAKKNLEVKKNCLILUSFSEA 378

## RESULT 5

US-09-407-367-6  
 ; Sequence 6, Application US/09407367  
 ; Patent No. 6420532  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Christophe P.G. Gerald, et al.  
 ; TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2 SPECIFIC COM  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/407,367  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPW  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-391-0525  
 ; TELEFAX: 212-391-0525  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 381 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-407-367-6

Query Match 21.2%, Score 474; DR 4; Length 381;  
 Best Local Similarity 29.6%, Pred. No. 11e-32;  
 Matches 119; Conservative 78; Mismatches 145; Indels 60; Gaps 12;  
 QY 17 DHNLTREOFIALYRLRLVLTPELPG-----RAKIALVLTGVLIFALALPQN 63  
 Db 9 DENOTVEKVFYSGGTPTRGELPPDPELIDSTKLVEQVVLILAYCSILLGVGN 68  
 QY 64 ALVFVWVTRSKAMRTVNIIFCSLALSDDLITFCIPVTMLQNSDNLGGAFCIKMVPF 123  
 Db 69 SLVTHVVIKPKSMVIVNFFIANLAVADLLVNLICLPTLTITLTMGEKMGCVLCHLVY 128  
 QY 124 VOSTAVVTEILTMTCIAVERHQGLVHPFKMKQVYNNRAFTMLGVVWLVAIVGSPMHW 183  
 Db 129 AGGLAVQVSTITLVIALDRHRCIVYHLESK--ISKQISFLLIIGLAWGVSALLASPLAIF 186  
 QY 184 QO---LEIKYDFLYEKHEHCCLPEW---TSPVQKQIYTTTFLVILPLMWMILYLSKI 237  
 Db 187 REYSLIEIIPDF---EIVACTEKWGEKSVYGVTSLSLTLLVYLPGLGISFSYTRI 242  
 QY 238 GYELW--IKKRVGDSVLETHGKEMSKIAKKKKRAVINMMVTVALFVAGWAFPHVIMM 295  
 Db 243 ---WSKLNKHNVSFCAASDHYH-----ORRHKKTKMLVCVVVFAVSWLPLHAFOLA 290

QY 296 IYNSNFE---KEYDDVTIKMIFALVIGFNSNSICNPVIVYAPMNFKNKNSVAVCYCIV 352  
 Db 291 VDISHVLDELKEY---KLITFVFIHIAWCSTEANPLCYGMNSNYKAFLSA----- 339  
 QY 354 NKTFSPAQR---HGNSGITMMKKAKFSLRNPVEETKGEA 390  
 Db 340 -FRCHQRLDATHSEVSMTFKAKKNIYVKKKNGLTDSFSEA 378

## RESULT 6

US-08-192-288-2  
 : Sequence 2, Application US/08192288  
 : Patent No. 5545549  
 : GENERAL INFORMATION:  
 : APPLICANT: Gerald, Christophe  
 : APPLICANT: Walker, Mary  
 : APPLICANT: Brancheck, Theresa  
 : APPLICANT: Weinshank, Richard L.  
 : TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE  
 : TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTOR AND USES  
 : NUMBER OF SEQUENCES: 2  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cooper & Dunham LLP  
 : STREET: 40 Rockefeller Plaza  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: U.S.A.  
 : ZIP: 10112

: COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent In Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/192,288  
 : FILING DATE:

: CLASSIFICATION: 514  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: White, John P.  
 : REGISTRATION NUMBER: 28,678  
 : REFERENCE/DOCKET NUMBER: 44742/JPW/TEP  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (212) 977-9550  
 : TELEFAX: (212) 664 0525  
 : TELEX: 422525 COOP UI  
 : INFORMATION FOR SEQ ID NO: 2:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 381 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-192-288-2

Query Match 21.0%; Score 469.5; DB 1; Length 381;  
 Best Local Similarity 31.0%; Pred. No. 2,60-32;  
 Matches 113; Conservative 75; Mismatches 123; Indels 53; Gaps 13;

QY 38 PELPGRALK-----ALVLTGVLIFALALFGLNALFYVYVTRSKAMRTVTNIFICSLASDL 92  
 Db 38 PELIDSTKLIHQVQVLLIAYGSIILGLVIGNSLVHVVVKPKSKMTVTNFFIANIAVADL 97  
 QY 93 LITFCIPVTMIQNTSDNLGAFICKMVPFQSTAVVTEILLTMTCTAVERHOGLVHPEK 152  
 Db 98 LVNTLCPLPTLYTLMGKWKMGVPLCHLVPAQGLAVQVSTITLTVALDRHRCIVYHLE 157  
 QY 153 MKWQYTNKRAFTMIGVWMLVAVIVGSPMHWVQO---LEIKYDFLYEKHEICCLPEW---T 206  
 Db 158 SK--ISKRSFLIIGLAWGISALLASPLAIFREYSLEIIPDF-----EIVACTEKWPGE 211  
 QY 207 SPVHUKIYTTFTLVILFLLPLMWMLILYSKTYGYELW---IKRPVGDGSLVRLTIHGKEMSKI 264  
 Db 212 KSIYGVTVSLSLLILYVLPILGIIISFSYTRI-----WSKLKNHVSFGAANDHYH----- 260

QY 265 ARKKRAVIMVTVVALFVACWAPRHHVMMIYSNFE---KEYDDVTIKMIFALVIG 321  
 Db 261 -QRRQKTTKMLCVVVVFAVSWLPLAFOLAVDIDSQVLDLKEY-----KLITFVFIHIA 314  
 QY 322 FSNISICNPVIVYAFMNFKNKNSVAVCYCIVNKTFSPAQR---HGNSGITMMKKAKFS 377  
 Db 315 MCSTFANPLIYCYMNSNYKAFLSA-----FRCHQRLDATHSEVSVTFKAKK NLF 364  
 QY 378 LREN 381  
 Db 365 VRKN 368

## RESULT 7

US-08-687-355A-2  
 : Sequence 2, Application US/08687355A  
 : Patent No. 5989834  
 : GENERAL INFORMATION:  
 : APPLICANT: Synaptic Pharmaceutical Corporation  
 : TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE  
 : TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF  
 : NUMBER OF SEQUENCES: 27  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cooper & Dunham LLP  
 : STREET: 1185 Avenue of the Americas  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: U.S.A.  
 : ZIP: 10036

: COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent In Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/687,355A  
 : FILING DATE: NO. 5989834ember 26, 1996  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: White, John P.  
 : REGISTRATION NUMBER: 28,678  
 : REFERENCE/DOCKET NUMBER: 44742-A-PCI/JPW/MAT  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 212-278-0400  
 : TELEFAX: 212-391-0525  
 : INFORMATION FOR SEQ ID NO: 2:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 381 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-687-355A-2

Query Match 21.0%; Score 469.5; DB 2; Length 381;  
 Best Local Similarity 31.0%; Pred. No. 2,60-32;  
 Matches 113; Conservative 75; Mismatches 123; Indels 53; Gaps 13;

QY 38 PELPGRALK-----ALVLTGVLIFALALFGLNALFYVYVTRSKAMRTVTNIFICSLASDL 92  
 Db 38 PELIDSTKLIHQVQVLLIAYGSIILGLVIGNSLVHVVVKPKSKMTVTNFFIANIAVADL 97  
 QY 93 LITFCIPVTMIQNTSDNLGAFICKMVPFQSTAVVTEILLTMTCTAVERHOGLVHPEK 152  
 Db 98 LVNTLCPLPTLYTLMGKWKMGVPLCHLVPAQGLAVQVSTITLTVALDRHRCIVYHLE 157  
 QY 153 MKWQYTNKRAFTMIGVWMLVAVIVGSPMHWVQO---LEIKYDFLYEKHEICCLPEW---T 206  
 Db 158 SK--ISKRSFLIIGLAWGISALLASPLAIFREYSLEIIPDF-----EIVACTEKWPGE 211  
 QY 207 SPVHUKIYTTFTLVILFLLPLMWMLILYSKTYGYELW---IKRPVGDGSLVRLTIHGKEMSKI 264  
 Db 212 KSIYGVTVSLSLLILYVLPILGIIISFSYTRI-----WSKLKNHVSFGAANDHYH----- 260



```

RESULT 10
US-08-876-798A-2
; Sequence 2, Application US/08876798A
; Patent No. 6355478
; GENERAL INFORMATION:
; APPLICANT: Baez, Melvyn
; APPLICANT: Yang, Peiyi
; TITLE OF INVENTION: RHESUS MONKEY NEUROPEPTIDE Y RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/876,798A
; FILING DATE: 16-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 46,808
; REFERENCE/DOCKET NUMBER: X 10901
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)276-0756
; TELEFAX: (317)276-2763
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-876-798A-2

```

```

Query Match 21.0%; Score 468.5; DB 4; Length 381;
Best Local Similarity 30.8%; Pred. No. 3,20-32;
Matches 112; Conservative 76; Mismatches 123; Indels 53; Gaps 13;

QY 38 PELPGRAKLI-----ALVLTGVLIFALALFGNALVFYVYVTRSKAMRTVTNIFICSLASDL 92
DB 38 PELIDSKLIEVQVVLILAYCSIIILGVIGNSLIHVVYIKFSMRTVTINEFIANLAVADL 97
QY 93 LITFFCIPVTMIQNIISNWLGGAFICKMVPVQSTAVVTEIITMTCIAVERHOGILVHPK 152
DB 98 VVNTLCPLPTLTLYTLMGFWKKMPVLCHLVPAQGLAVQVSTITLTVALDHRHCIVVILE 157
QY 153 HKQYTNRAFTMLGVVWLVAVVGVSPMHIQQ---LEIKYDFLYEKEHICGLEW---T 206
DB 158 SK-ISKRIISFLIIGLAWCISALLASPLAIFREYSLEIIPDF----EIVACTPKMGEE 211
QY 207 SPVHKIYTTFTIIVLFLPLMVLILYSKIGYELW--IKKPGVGVSVLPTIHKCKSKI 254
DB 212 KSIYTVLSLSLLIYVPLGITSFYTRI---WSKLKSHVSPGAANDHYH-----260
QY 265 ARKKKRAVIMVTVVAVFVAVVAVVHVHMMIYSNFE---KEYDDVTIKMIFAIVQIG 321
DB 261 -QRRUKTKMLGVVWVFAVSWLPHAFQAVLDIDSHVLDLKEY-----KLITVFIHIA 314
QY 322 FSNISICNPVIAFMNENPKKVLNAVYGVYVKNRTFSPAQR-----HGNSGIMTKRKAKFS 377
DB 315 MCFSTFANPLLYCWMNSNYRKAFUSA-----FRCEQRJDAIHSEVSVTFKAKK-NLE 364
QY 478 LREN 481
DB 465 VRKN 468

```

```

RESULT 11
US-09-426-290-2
; Sequence 2, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Herglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345,2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-426-290 2

Query Match 20.4%; Score 455; DB 4; Length 444;
Best Local Similarity 27.7%; Pred. No. 5,3e-31;
Matches 104; Conservative 91; Mismatches 142; Indels 38; Gaps 9;

QY 4 LNIPTQFSRLLRHNLTRHQFIAYLRPLVYTPPELPGRAKIALVITGVLIIFALIPGN 63
DB 21 LNETQPPF--INPTDYDDDEFLR-YLWREYLH----PKYEWVLIAGYIIVFWVALIGN 72
QY 64 ALVEYVTVTRSKAMRTVTNIFICSLASDLITFFCIPVTMIQNIISNWLGGAFICKMVP 123
DB 73 VLYCVAVVKNHMMRTVTNYFIVNISLADVLVITCLIPATIVVDITETWFFQSLCKVDPY 132
QY 124 VOSTAVVTEIITMTCIAVERHOGILVHPKMKMVOYTNRAFTMIQVWLVAVVGVSPMHHV 183
DB 133 LOTSVSVSVLTLSCLALDRWYAICHPLMEK--STAKRARNSTIVITWVSCIIMIPQAI 190
QY 184 QOLEIKYDFLYEKEHIL--CCLEFMTSPVHOKIYTTFTIIVLFLPLMVLILYSKIGYEL 241
DB 191 MECSTVFPCLANKTTITFTVCDERWGGELIYKMYHICFPIVYTMAPICLMVLAVIQIFRKL 250
QY 242 WTKRVRGDSVLR-----TIHGKMSKIARKKKRAVIMVTV 278
DB 251 WCKQIPLGLISSVQKKWKPLQVSPQPGHPTKSRMSAAAEIKQIRARKKTARMLMV-V 309
QY 279 VALFVAVVAVVHVHMMIE-YSNFEKYDDVTIKMIFAIVQIIGFNSICNPVIAFMNE 337
DB 310 LLVFAICLPLISILNVLKRVGMEFAHTEDETVAVFTFSHMLVYANSAANPIIYNELSG 369
QY 338 NFKNVLSAV-CYCI 351
DB 370 KPRFEKAAFPSCCCL 384

```

```

RESULT 12
US-09-119-788-2
; Sequence 2, Application US/09119788
; Patent No. 6166193
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; TITLE OF INVENTION: CNA CLONE MY1 THAT ENCODES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,788
; FILING DATE: 21-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/053,790
; FILING DATE: 25-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T.
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GH50029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5515
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 444 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-119-788-2

```

Query Match 20.3%; Score 453; DB 4; Length 444;

Best Local Similarity 27.7%; Pred. No. 7.8e-31;

Matches 104; Conservative 90; Mismatches 143; Indels 38; Gaps 9;

```

QY 4 LNITPEQFSLRDHNLTRQOFIALYRLPLVYTPELGRKALALVLTGVLIFALFALPGN 63
DB 21 LNITQEPF---LNPTDYDDEFLR-YLWREYLDH---PKREYEWLIAGYIIVFVALIGN 72
QY 64 ALVYVYVTRSKAMPTVNTIFCSIALSDLLITFFCIPVPTMLOINISNNLGGAFICKMVPF 123
DB 73 VLVCVAVKNNHMKRTVNTYFVNLSDVLVLTICLPATLVVDITETWTFEGOSLCKVIPY 132
QY 124 VOSTAVVTELTMTCTAVERHOGVLVHPFKMKWOYTNNRAFTMLGVVWLVAIVGSPMHIV 183
DB 133 LQTVSVSVSLTISALDKRWYAICHPLMEK---STAKRARNISIVITWVSCIIMIPQALV 190
QY 184 QOLEIKYDFLYEKEHI---CCLEWTSVPVHOKIYTTFFILVILFLPLMVMLILYSKIGYEL 241
DB 191 MEGSTVEPGLANKTTLFTVCDERWGEIYPKMYHICFFLVTYNAPLGLMLVLAQLQIFPKI 250
QY 242 WIKKKVGDGSLVLT-----IHGKEMSKIAKKKRAVIMMVTV 278
DB 251 WCROIPTGSSWQKKKPLQPVSGPRGPGOPTKSRMGAAVAETKQIRARKKTARMIMV-V 309
QY 279 VALFVACWAPFHVHMMIE-YSNFEKEYDDVTIKMIFAIVQIGFSNSICNPIVYAPNNE 337
DB 310 LJVFAICVTPISLIINVLKRVFGMPAHTPEFTVYAMFTFSHWIIVVANSAAANPIYNHLSG 359
QY 338 NFKKNVLSAV-CYCI 351
DB 370 KREEFKAAFSCCL 384

```

#### RESULT 13

US-09-255-368-2

; Sequence 2, Application US/09255368

; Patent No. 6262246

; GENERAL INFORMATION:

; APPLICANT: Gerald, Christophe P.G.

; APPLICANT: Jones, Kenneth A.

; APPLICANT: Bonini, James A.

; APPLICANT: Borowsky, Beth

; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors

; TITLE OF INVENTION: and Uses Thereof

; FILE REFERENCE: 1795/57155-A

; CURRENT APPLICATION NUMBER: US/09/255,368

; CURRENT FILING DATE: 1999-02-22

; EARLIER APPLICATION NUMBER: 09/161,113

; EARLIER FILING DATE: 1998-09-25

```

; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 2
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-255-368-2

```

Query Match 19.6%; Score 437; DB 4; Length 432;

Best Local Similarity 28.1%; Pred. No. 1.7e-29;

Matches 101; Conservative 70; Mismatches 119; Indels 70; Gaps 8;

```

QY 11 FSRLLRDHNLTRQOFIALYRLPLVYTPELGRKALALVLTGVLIFALFALFALVAVV 70
DB 32 FSSYYQHSSPVAAAFIAAY-----VLIFLLCMVGNLTIVGPIV 68
QY 71 TRSKAMRTVNTIFCSIALSDLLITFFCIPVPTMLOINISNNLGGAFICKMVPFVQSTAVV 130
DB 69 LKNRHMPVTNMFILNLAIVSDLLVGFICMPTTLVDNLITGWPEDNATCKMSLSLVGMSVS 128
QY 131 TEILTMTCTAVERHOGVLVHPFKMKWOYTNNRAFTMLGVVWLVAIVGSPMHVVOOLEIKY 190
DB 129 ASVPTIIVAJAVERFCIVHPFREK--LTURKALFTIAVIALALLIMCPS-----AVTL 180
QY 191 DFLYEKEHI-----CCLEWTSVPVHOKIYTTFFILVILFLPLMVMLILYSKIG 238
DB 181 TVTTRHHFMDARNRSYPLYSQFAPWPEKGRKVTAVLFAHIVLPLALIVVYVRIA 240
QY 239 YELW-----IKRVGDSGLRTIHGKEMSKIAKKKRAVIMMVTVVAVFACWAPFH 290
DB 241 RKLQAPGAPADTEEAAGE-----GTPSPRPAPVVMILVMVALFETLSWLPW 289
QY 291 VVHMMIEYSNFEKEYDDVTIKMI---FAIVQIGFSNSICNPIVYAFMNNKKNVLSA 346
DB 290 VLLLLLIYDG---ELSELQLHLLSVYAFPLAHWIAFFHSSANPIIYGFNENFRGQAA 345

```

#### RESULT 14

US-09-255-368-8

; Sequence 8, Application US/09255368

; Patent No. 6262246

; GENERAL INFORMATION:

; APPLICANT: Gerald, Christophe P.G.

; APPLICANT: Jones, Kenneth A.

; APPLICANT: Bonini, James A.

; APPLICANT: Borowsky, Beth

; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors

; TITLE OF INVENTION: and Uses Thereof

; FILE REFERENCE: 1795/57155-A

; CURRENT APPLICATION NUMBER: US/09/255,368

; CURRENT FILING DATE: 1999-02-22

; EARLIER APPLICATION NUMBER: 09/161,113

; EARLIER FILING DATE: 1998-09-25

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn Ver. 2.0 - beta

; SEQ ID NO 8

; LENGTH: 430

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-255-368-8

Query Match

Best Local Similarity 29.8%; Pred. No. 4.1e-29;

Matches 93; Conservative 64; Mismatches 126; Indels 29; Gaps 5;

```

QY 49 VLTGVLIFALFALFALVAVVYVTRSKAMRTVNTIFCSIALSDLLITFFCIPVPTMLOINIS 108
DB 47 IVAVALIFLCWGNLTIVGFIVLKRNHMTVNTMILNLAIVSDLLVGFICMPTTLVDNLI 106
QY 109 LNNIAGGAFICKMVPFVQSTAVVTEILTMTCTAVERHOGVLVHPFKMKWOYTNNRAFTMIGV 168
DB 107 TGMFPDNATCKMSGLVCGMSVSASVPTLVAIAVEFPIVHPFREK--LTLPKALVITAV 164

```

```

QY 169 VMLVAVVIGSPMHVQQLIKYDFLYEKEHI-----CCLEFWTSPVHOKIYTT 216
Db 165 IVALALLIMCP-----AVLTVTREEHVWVARNRSYPLYSWEAWPEKGMRYTT 218
QY 217 FILVILFLPLVMVLLSKIGYELWIKKPVGVGVSVLRTINGKMS--KIAPKKKRAVIM 274
Db 219 VLFSHIYIAPLALIVVMVARIARKLCQAPGAPG-----GEEAADPRASRRARVVMH 271
QY 275 MVTVALFAVCWAPFVHMIMIEYSNFEKEYDDVTIKMIFAIVQIIGFSNCPVYAF 334
Db 272 LVNVALFTLSMLPLWALLLDYQLSAPQLHLVTYVAFPAHMLAFAFNSSANPLIYGY 331
QY 335 MNENFKKKNVLSA 346
Db 332 FNEFERRGFOAA 343

```

```

RESULT 15
US-09-199-737-5
: Sequence 5, Application US/09199747A
: Patent No. 6287788
: GENERAL INFORMATION:
: APPLICANT: Bard, Jonathan A.
: APPLICANT: Borowsky, Beth
: APPLICANT: Smith, Kelli E.
: APPLICANT: Branchek, Theresa A.
: APPLICANT: Gerald, Christophe P.G.
: APPLICANT: Jones, Kenneth A.
: TITLE OF INVENTION: DNA Encoding Galanin GALR3 Receptors And Uses Thereof
: FILE REFERENCE: 52241-D-PCT-US
: CURRENT APPLICATION NUMBER: US/09/199,737A
: NUMBER OF SEQ ID NOS: 59
: SOFTWARE: Patent In Ver. 2.0 - beta
: SEQ ID NO 5
: LENGTH: 346
: TYPE: PRT
: ORGANISM: Rat
US-09-199-737-5

```

```

Query Match 19.38; Score 431.5; DB 4; Length 346;
Best Local Similarity 32.18; Pred. No. 3,9e-29;
Matches 103; Conservative 60; Mismatches 115; Indels 43; Gaps 10;

QY 48 PELGRAK-----LALVLTGVLIFALFGLNALVYVYVTRSK--AMRTVTNIFIC 85
Db 16 PEPPAEPRPLRGVIGVNFITILWVG-LIFAMGVLSNLSLITVLARSKPGKPRSTNLFIL 74
QY 86 SLALSGLLTFFCIHVIMIONISNWLGGAFICKMVPFVOSTAVVTKILMTCLAVRHO 145
Db 75 NLSIADLAYLLFCIPFQATVVALPTWLGAFICKFIIHYFTVSMVLSIFTLAAMSVDYV 134
QY 146 GLVHPFKMKWYTNRAFTMLGVVNLVAVVIGSPMHVQQLIKYDFLYEKEHICLEEW 205
Db 135 ALVHSRRSSSLRVSNALLGVGFIALSLAMASHVAYYQRL-----PHRDSNOTFCWEHW 189
QY 206 TSPVIOKIYTTFILVILFLPLVMVLLSKIGYELWIKKPVGVGVSVLRTING--KEMSK 263
Db 190 PQLHKKAYVVTVEGVLPLLELICFCYAK-----VLNHLHKKLNMSK 234
QY 264--IARKKKRAVIMVTVVALFAVCWAPFVHMIMIEYSNFEKEYDDVTIKMIFAIVQIIGFSNCPVYAF 334
Db 245 KSEASKKKTAUTVLV-VVVVRGISKMLPHHVIHLWAERCAAPL-----TPASFFRITAIKCL 289
QY 321 GFSNSICNPVYAFMNFKK 341
Db 240 AYSNSSVNPVYAFISENFK 310

```

Search completed: June 10, 2003, 14:18:47  
Job time : 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: Tue 10, 2003, 13:17:03, Search time 25 seconds  
(without alignments)  
1779 862 Million cell updates/sec

Title: US-10-070-241R-1

Perfect score: 2235

Sequence: 1 MQALNITPEQFSLPDHNL RHILFRSELAENSLDSH 431

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 392085 seqs, 101240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep:\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep:\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep:\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep:\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep:\*

6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep:\*

7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep:\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep:\*

9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep:\*

10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep:\*

11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep:\*

12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*

13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep:\*

14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description       |
|------------|-------|---------------|--------|-------|-------------------|
| 1          | 2235  | 100.0         | 431    | 9     | US-09-966-782A-10 |
| 2          | 2235  | 100.0         | 431    | 10    | US-09-966-782A-11 |
| 3          | 2235  | 100.0         | 431    | 10    | US-09-966-782A-12 |
| 4          | 1914  | 85.6          | 433    | 9     | US-09-966-782A-16 |
| 5          | 1767  | 79.1          | 416    | 9     | US-09-966-782A-18 |
| 6          | 502   | 22.5          | 420    | 9     | US-09-966-782A-6  |
| 7          | 479   | 21.4          | 381    | 9     | US-10-067-649-5   |
| 8          | 469   | 21.0          | 381    | 10    | US-09-966-782A-16 |
| 9          | 469   | 21.0          | 381    | 10    | US-09-966-782A-17 |
| 10         | 469.5 | 21.0          | 381    | 10    | US-09-966-782A-20 |
| 11         | 466.5 | 20.9          | 385    | 10    | US-09-966-782A-18 |
| 12         | 455   | 20.4          | 444    | 9     | US-09-966-782A-19 |
| 13         | 453   | 20.3          | 444    | 9     | US-10-067-649-6   |
| 14         | 440   | 19.7          | 385    | 9     | US-10-067-649-6   |
| 15         | 437   | 19.6          | 432    | 9     | US-09-966-782A-2  |
| 16         | 434.5 | 19.4          | 348    | 9     | US-09-966-782A-9  |
| 17         | 432.5 | 19.4          | 428    | 10    | US-09-966-782A-4  |
| 18         | 432.5 | 19.4          | 430    | 9     | US-09-966-782A-8  |
| 19         | 431.5 | 19.3          | 346    | 1     | US-08-899-112-32  |

Sequence 10, Appl  
Sequence 11, Appl  
Sequence 6, Appl  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 2, Appl  
Sequence 10, Appl  
Sequence 5, Appl  
Sequence 17, Appl  
Sequence 3, Appl  
Sequence 19, Appl  
Sequence 2, Appl  
Sequence 15, Appl  
Sequence 24, Appl  
Sequence 22, Appl  
Sequence 19, Appl  
Sequence 23, Appl  
Sequence 4, Appl  
Sequence 26, Appl  
Sequence 2, Appl  
Sequence 5, Appl  
Sequence 12, Appl  
Sequence 7, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-966-782A-10  
; Sequence 6, Application US/09990940  
; Publication No. US20030027252A1  
; GENERAL INFORMATION:  
; APPLICANT: Tian, Hui  
; APPLICANT: Zhao, Jiaqiang  
; APPLICANT: Chen, Jin-Long  
; APPLICANT: Catlet, Gene  
; APPLICANT: An, Songzhu  
; APPLICANT: Dai, Kang  
; APPLICANT: Gupta, Jamila S.  
; APPLICANT: Tularik Inc.  
; TITLE OF INVENTION: No. US20030027252A1c1 Receptors  
; FILE REFERENCE: US781-007410US  
; CURRENT APPLICATION NUMBER: US/09/990, 940  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/252,841  
; PRIOR FILING DATE: 2000-11-22  
; PRIOR APPLICATION NUMBER: US 60/257,636  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 60/261,377  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: US 60/279,554  
; PRIOR FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: US 60/280,696  
; PRIOR FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: FASTA2.1; Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR45  
US-09-966-782A-10

Query Match: 100.0%, Score 2235, 6B 9, Length 431;  
Best Local Similarity: 100.0%; Pred. No. 1.1e-195;  
Matches 431, Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQFSLPDHNLRTPEJFIALYFLHLYVTPKLPKPAKALVLTGLFALAL 60

```

Db      1  MOALNITPEQSRLLRDHNLTRQFIALYRLRPLVYTPPELPGRAKALALVLTGVLIFALAL 60
QY      61  FGNALVYVYVTRSKAMRTVTNIFIGSLALSDDLITFFCIPVTMLONISDNWLGGAFCCKM 120
Db      61  FGNALVYVYVTRSKAMRTVTNIFIGSLALSDDLITFFCIPVTMLONISDNWLGGAFCCKM 120
QY      121  VPFVQSTAVVTEILTMTCTIAVERHUGLVHPFKMKQYTNRRAFITMLGVVWLVAIVGSPM 180
Db      121  VPFVQSTAVVTEILTMTCTIAVERHUGLVHPFKMKQYTNRRAFITMLGVVWLVAIVGSPM 180
QY      181  WHVQOOLEIKYDFLYEKHEHICCELEWTSPVHQKIYTTFLVILPLVMWLLIYSKIGYE 240
Db      181  WHVQOOLEIKYDFLYEKHEHICCELEWTSPVHQKIYTTFLVILPLVMWLLIYSKIGYE 240
QY      241  LWIKKRVGDSVLRTHCKEMSKIAKKKRAVIMMVTVAALFVCAWAPFHVHMMIEYSN 300
Db      241  LWIKKRVGDSVLRTHCKEMSKIAKKKRAVIMMVTVAALFVCAWAPFHVHMMIEYSN 300
QY      301  FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFKNKVL SAVCYCIVNKTSPAQ 360
Db      301  FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFKNKVL SAVCYCIVNKTSPAQ 360
QY      361  RHGNSGITMMRKAKFSLRENIVERTKGEAFSDGNIEVKLCQOTPEKKKIKRHIALPRSE 420
Db      361  RHGNSGITMMRKAKFSLRENIVERTKGEAFSDGNIEVKLCQOTPEKKKIKRHIALPRSE 420
QY      421  LAENSPDLSGH 431
Db      421  LAENSPDLSGH 431

```

## RESULT 2

```

US-09-826-508 22
; Sequence 22, Application US/09826508
; Patent No. US2001002509A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbary
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; FILE REFERENCE: GP 70744US
; CURRENT APPLICATION NUMBER: US/09/826,508
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 431
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-826-508 22

```

```

Query Match
Best Local Similarity 100.0%; Score 2235; DB 10; Length 431;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MOALNITPEQSRLLRDHNLTRQFIALYRLRPLVYTPPELPGRAKALALVLTGVLIFALAL 60
Db      1  MOALNITPEQSRLLRDHNLTRQFIALYRLRPLVYTPPELPGRAKALALVLTGVLIFALAL 60
QY      61  FGNALVYVYVTRSKAMRTVTNIFIGSLALSDDLITFFCIPVTMLONISDNWLGGAFCCKM 120
Db      61  FGNALVYVYVTRSKAMRTVTNIFIGSLALSDDLITFFCIPVTMLONISDNWLGGAFCCKM 120
QY      121  VPFVQSTAVVTEILTMTCTIAVERHUGLVHPFKMKQYTNRRAFITMLGVVWLVAIVGSPM 180
Db      121  VPFVQSTAVVTEILTMTCTIAVERHUGLVHPFKMKQYTNRRAFITMLGVVWLVAIVGSPM 180
QY      181  WHVQOOLEIKYDFLYEKHEHICCELEWTSPVHQKIYTTFLVILPLVMWLLIYSKIGYE 240
Db      181  WHVQOOLEIKYDFLYEKHEHICCELEWTSPVHQKIYTTFLVILPLVMWLLIYSKIGYE 240
QY      241  LWIKKRVGDSVLRTHCKEMSKIAKKKRAVIMMVTVAALFVCAWAPFHVHMMIEYSN 300

```

```

Db      241  LWIKKRVGDSVLRTHCKEMSKIAKKKRAVIMMVTVAALFVCAWAPFHVHMMIEYSN 300
QY      301  FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFKNKVL SAVCYCIVNKTSPAQ 360
Db      301  FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFKNKVL SAVCYCIVNKTSPAQ 360
QY      361  RHGNSGITMMRKAKFSLRENIVERTKGEAFSDGNIEVKLCQOTPEKKKIKRHIALPRSE 420
Db      361  RHGNSGITMMRKAKFSLRENIVERTKGEAFSDGNIEVKLCQOTPEKKKIKRHIALPRSE 420
QY      421  LAENSPDLSGH 431
Db      421  LAENSPDLSGH 431

RESULT 3
US-09-899-532-2
; Sequence 2, Application US/09899532
; Patent No. US20020048791A1
; GENERAL INFORMATION:
; APPLICANT: Bloomquist, Brian T.
; APPLICANT: Zhebin, Leonid
; TITLE OF INVENTION: Human Neuroreceptor Y like G
; TITLE OF INVENTION: Protein-Coupled Receptor
; FILE REFERENCE: 02973.00040
; CURRENT APPLICATION NUMBER: US/09/899,532
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/216,523
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-532-2

```

```

Query Match
Best Local Similarity 100.0%; Score 2235; DB 10; Length 431;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MOALNITPEQSRLLRDHNLTRQFIALYRLRPLVYTPPELPGRAKALALVLTGVLIFALAL 60
Db      1  MOALNITPEQSRLLRDHNLTRQFIALYRLRPLVYTPPELPGRAKALALVLTGVLIFALAL 60
QY      61  FGNALVYVYVTRSKAMRTVTNIFIGSLALSDDLITFFCIPVTMLONISDNWLGGAFCCKM 120
Db      61  FGNALVYVYVTRSKAMRTVTNIFIGSLALSDDLITFFCIPVTMLONISDNWLGGAFCCKM 120
QY      121  VPFVQSTAVVTEILTMTCTIAVERHUGLVHPFKMKQYTNRRAFITMLGVVWLVAIVGSPM 180
Db      121  VPFVQSTAVVTEILTMTCTIAVERHUGLVHPFKMKQYTNRRAFITMLGVVWLVAIVGSPM 180
QY      181  WHVQOOLEIKYDFLYEKHEHICCELEWTSPVHQKIYTTFLVILPLVMWLLIYSKIGYE 240
Db      181  WHVQOOLEIKYDFLYEKHEHICCELEWTSPVHQKIYTTFLVILPLVMWLLIYSKIGYE 240
QY      241  LWIKKRVGDSVLRTHCKEMSKIAKKKRAVIMMVTVAALFVCAWAPFHVHMMIEYSN 300
Db      241  LWIKKRVGDSVLRTHCKEMSKIAKKKRAVIMMVTVAALFVCAWAPFHVHMMIEYSN 300
QY      301  FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFKNKVL SAVCYCIVNKTSPAQ 360
Db      301  FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFKNKVL SAVCYCIVNKTSPAQ 360
QY      361  RHGNSGITMMRKAKFSLRENIVERTKGEAFSDGNIEVKLCQOTPEKKKIKRHIALPRSE 420
Db      361  RHGNSGITMMRKAKFSLRENIVERTKGEAFSDGNIEVKLCQOTPEKKKIKRHIALPRSE 420
QY      421  LAENSPDLSGH 431
Db      421  LAENSPDLSGH 431

```



## RESULT 4

US-09-990-940-16  
 ; Sequence 16, Application US/09990940  
 ; Publication No. US20030027252A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tian, Hui  
 ; APPLICANT: Zhao, Jiagang  
 ; APPLICANT: Chen, Jin-long  
 ; APPLICANT: Cutler, Gene  
 ; APPLICANT: An, Songzhu  
 ; APPLICANT: Dai, Kang  
 ; APPLICANT: Gupta, Jamila S.  
 ; APPLICANT: Tularik Inc.  
 ; TITLE OF INVENTION: NO. US20030027252A1c1 Receptors  
 ; FILE REFERENCE: 018781-0074100S  
 ; CURRENT APPLICATION NUMBER: US/09/990, 940  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: US 60/252, 841  
 ; PRIOR FILING DATE: 2000-11-22  
 ; PRIOR APPLICATION NUMBER: US 60/257, 636  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: US 60/261, 377  
 ; PRIOR FILING DATE: 2001-01-12  
 ; PRIOR APPLICATION NUMBER: US 60/279, 554  
 ; PRIOR FILING DATE: 2001-03-28  
 ; PRIOR APPLICATION NUMBER: US 60/280, 696  
 ; PRIOR FILING DATE: 2001-03-29  
 ; NUMBER OF SEQ ID NOS: 54  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 16  
 ; LENGTH: 433  
 ; TYPE: PRT  
 ; ORGANISM: Mus sp.  
 ; FEATURE:  
 ; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR346a  
 US-09-990-940-16

Query Match 85.6%; Score 1914; DB 9; Length 433;  
 Best Local Similarity 83.5%; Pred No. 2.1e-166;  
 Matches 360; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

QY 1 MOALNITPEQFSRLRLDHNLTREQFIATYRLRLPLVTPPELPGRAKIALVLTGVLPALAL 60  
 DB 1 MOALNITAEQFSRLLSAHLNLTREQFIHYGLRLPLVTPPELPGRAKIALVLTGVLPALAL 60

QY 61 FGNALVYVYVTRSKAMRTVNTIFICSLSLSDLLITTEFCIPVTMLQNSDNWLGGAFCIKM 120  
 DB 61 FGNLSLVYVYVTRSKAMRTVNTIFICSLSLSDLLITTEFCIPVTMLQNSDKWLGGAFCIKM 120

QY 121 VPVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRFTMTLGVVWLVAVIVGSPM 180  
 DB 121 VPVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRFTMTLGVVWLVAAIVGSPM 180

QY 181 WHVQGLEIKDYFEKEKHICCLEEWSVPHQKIYTTFTLLVILPLPLVWMLVLYSKIGYE 240  
 DB 181 WHVQGLEIKDYFEKEKHICCLEEWSVPHQKIYTTFTLLVILPLPLVWMLVLYSKIGYE 240

QY 241 LWIKKRGDGSVLTTHGKEMSKTARKKRAVIMVTVVVALFAVWAPFVHVHMMIEYSN 300  
 DB 241 LWIKKRGDGSVLTTHGKEMSKTARKKRAVIMVTVVVALFAVWAPFVHVHMMIEYSN 300

QY 301 FEKEYDDVTIKMIFAVQIIGFSNSICNPVYAFMNFENKKNVLSAVCYCIVNKTFSPAQ 360  
 DB 301 FEKEYDDVTIKMIFAVQIIGFSNSICNPVYAFMNFENKKNVLSAVCYCIVNKTFSPOQ 360

QY 361 RHNSGTMKAKKAKFSLPENPVEFTKGEAFSCNTEVKECEOTEKELKRLIALFRSE 420  
 DB 361 KPGNSGTMKAKKAKFSLPENPVEFTKGEAFSCNTEVKECEOTEKELKRLIALFRSE 420

QY 421 LAENSPDLSGH 431  
 DB 421 LAENSPDLSGH 431

DB 421 LSENSTGSGH 431

## RESULT 5

US-09-990-940-18  
 ; Sequence 18, Application US/09990940  
 ; Publication No. US20030027252A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tian, Hui  
 ; APPLICANT: Zhao, Jiagang  
 ; APPLICANT: Chen, Jin-long  
 ; APPLICANT: Cutler, Gene  
 ; APPLICANT: An, Songzhu  
 ; APPLICANT: Dai, Kang  
 ; APPLICANT: Gupta, Jamila S.  
 ; APPLICANT: Tularik Inc.  
 ; TITLE OF INVENTION: NO. US20030027252A1c1 Receptors  
 ; FILE REFERENCE: 018781-0074100S  
 ; CURRENT APPLICATION NUMBER: US/09/990, 940  
 ; CURRENT FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: US 60/252, 841  
 ; PRIOR FILING DATE: 2000-11-22  
 ; PRIOR APPLICATION NUMBER: US 60/257, 636  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: US 60/261, 377  
 ; PRIOR FILING DATE: 2001-01-12  
 ; PRIOR APPLICATION NUMBER: US 60/279, 554  
 ; PRIOR FILING DATE: 2001-03-28  
 ; PRIOR APPLICATION NUMBER: US 60/280, 696  
 ; PRIOR FILING DATE: 2001-03-29  
 ; NUMBER OF SEQ ID NOS: 54  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 18  
 ; LENGTH: 416  
 ; TYPE: PRT  
 ; ORGANISM: Mus sp.  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(1302)  
 ; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR346b  
 US-09-990-940-18

Query Match 79.1%; Score 1767; DB 9; Length 416;  
 Best Local Similarity 80.1%; Pred No. 5.2e-153;  
 Matches 339; Conservative 36; Mismatches 38; Indels 10; Gaps 3;

QY 5 NITEPQFSRLRLDHNLTREQFIATYRLRLPLVTPPELPGRAKIALVLTGVLPALALPNA 64  
 DB 4 NLTAEQLSALLRGHNLTREQFIATYRLRLPLVTPPELPGRAKIALVLTGVLPALALPNA 63

QY 65 LVFYVYVYVTRSKAMRTVNTIFICSLSLSDLLITTEFCIPVTMLQNSDNWLGGAFCIKMVPV 124  
 DB 64 LVYVYVYVTRSKAMRTVNTIFICSLSLSDLLITTEFCIPVTMLQNSDNWLGGAFCIKMVPV 123

QY 125 QSTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRFTMTLGVVWLVAVIVGSPMHWQ 184  
 DB 124 QCTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRFTMTLGVVWLVAVIVGSPMHWQ 183

QY 185 QLETKYDPLFEKEKHICCLEEWSVPHQKIYTTFTLLVILPLPLVWMLVLYSKIGYELWK 244  
 DB 184 RLETKYDPLFEKEKHICCLEEWSVPHQKIYTTFTLLVILPLPLVWMLVLYSKIGYELWK 243

QY 245 KRVGDGSVLTTHGKEMSKTARKKRAVIMVTVVVALFAVWAPFVHVHMMIEYSNFKKE 304  
 DB 244 KRIGGVSVLTTHGKEMSKTARKKRAVIMVTVVVALFAVWAPFVHVHMMIEYSNFKKE 303

QY 305 YDDVTIKMIFAVQIIGFSNSICNPVYAFMNFENKKNVLSAVCYCIVNKTFSPAORIGN 364  
 DB 304 YDEVTIKMIFAVQIIGFSNSICNPVYAFMNFENKKNVLSAVCYCIVNKTFSPAORIGN 363

QY 365 SGITMMKKAFLSRLPENPVEFTKGEAFSCNTEVKECEOTEKELKRLIALFRSEIAFN 424  
 DB 364 SGAMVMPRAKLAARENPV-EIKGEAFGGSNDIDKWKCEQEKKR-----RSKVA-S 413

QY 425 SPL 427

Db 414 CPL 416

# RESULT 6

US-09-866-248A-6

; Sequence 6, Application US/09866248A

; Publication No. US20020198367A1

; GENERAL INFORMATION:

; APPLICANT: Gerald, Christophe P.G.

; APPLICANT: Jones, Kenneth A.

; APPLICANT: Bonini, James A.

; APPLICANT: Borowsky, Beth

; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors

; FILE REFERENCE: 1795/57155-A

; CURRENT FILING DATE: 2001-06-25

; PRIOR FILING DATE: 1998-09-25

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: Patent In Ver. 2.0 - beta

; SEQ ID NO 6

; LENGTH: 420

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-866-248A-6

Query Match 22.5%; Score 502; DB 9; Length 420;

Best Local Similarity 40.1%; Pred. No. 1.3e-37;

Matches 128; Conservative 87; Mismatches 162; Indels 48; Gaps 12;

QY 14 LLDHNLTRQFIALYRLRPLVYTPPELPGRAKIALVITGVLIIFALALFGNALPVYVYTRS 73

Db 14 LLDHNLTRQFIALYRLRPLVYTPPELPGRAKIALVITGVLIIFALALFGNALPVYVYTRS 73

QY 74 KAMRVNTNIFCSIALSDLLITFFCIPTVMTLQNSIDNMLGAGFICKMVPVQSTAVVTEI 133

Db 74 KAMRVNTNIFCSIALSDLLITFFCIPTVMTLQNSIDNMLGAGFICKMVPVQSTAVVTEI 133

QY 74 KAMRVNTNIFCSIALSDLLITFFCIPTVMTLQNSIDNMLGAGFICKMVPVQSTAVVTEI 133

Db 74 KAMRVNTNIFCSIALSDLLITFFCIPTVMTLQNSIDNMLGAGFICKMVPVQSTAVVTEI 133

QY 134 LMTCTIAVERHGLVHPEKMKQWYTNRAFTMLGVWLVAVIVGSP---MMHVQOLEIKY 190

Db 134 LMTCTIAVERHGLVHPEKMKQWYTNRAFTMLGVWLVAVIVGSP---MMHVQOLEIKY 190

QY 191 DFLYKPKH ---ICCLEWTSVPHOKIYTTFTLIVTLFLLPMLVMLLYSKIGYELWIKK 245

Db 191 DFLYKPKH ---ICCLEWTSVPHOKIYTTFTLIVTLFLLPMLVMLLYSKIGYELWIKK 245

QY 191 RVRLNSQKTSVYVW---RETWPQEMPKIYTTVLEANIYLAFLSLIVIMYGRIGISLF--- 247

Db 191 RVRLNSQKTSVYVW---RETWPQEMPKIYTTVLEANIYLAFLSLIVIMYGRIGISLF--- 247

QY 246 RVCDGSLVLTICKEMSK---IAHKKKRAVIMVTVVAFVAVCWAPPHVHHMYSNF 301

Db 246 RVCDGSLVLTICKEMSK---IAHKKKRAVIMVTVVAFVAVCWAPPHVHHMYSNF 301

QY 248 ---RAAVPT---GRNQKQWVSVSKKOKIKMLLIVALLPILSMPLWIMLSYADIL 302

Db 248 ---RAAVPT---GRNQKQWVSVSKKOKIKMLLIVALLPILSMPLWIMLSYADIL 302

QY 302 EKEYDQVTIKMIFAIQVIGFSNSICNPVYAFNENEFKNVLSAVCYGVNKTFSAPQR 361

Db 302 EKEYDQVTIKMIFAIQVIGFSNSICNPVYAFNENEFKNVLSAVCYGVNKTFSAPQR 361

QY 303 SPNEQLINIVYIPPAHVALFQNSVNPVYGFNENRPFQFQAFQIQCOKHAKPWEA 362

Db 303 SPNEQLINIVYIPPAHVALFQNSVNPVYGFNENRPFQFQAFQIQCOKHAKPWEA 362

QY 362 HCNSGITMMRKKAKSLRENVEETKGA-PSDGNIEVKLICHTERKKIKRHLALPRSE 420

Db 362 HCNSGITMMRKKAKSLRENVEETKGA-PSDGNIEVKLICHTERKKIKRHLALPRSE 420

QY 463 YA ---LRAKSHVILNTSNQVQESTFQNPGETLLYRKSAERKQOE---LVNME 410

Db 463 YA ---LRAKSHVILNTSNQVQESTFQNPGETLLYRKSAERKQOE---LVNME 410

QY 421 LAENS 425

Db 421 LAENS 415

# RESULT 7

US-10-067-649-5

; Sequence 5, Application US/10067649

; Publication No. US20040100057A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY14, RELATED T

; FILE REFERENCE: 00118 NP

; CURRENT APPLICATION NUMBER: US/10/067,649

; PRIOR FILING DATE: 2002-02-05

; PRIOR APPLICATION NUMBER: US 60/266,525

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: US 60/329,897

; PRIOR FILING DATE: 2001-10-16

; NUMBER OF SEQ ID NOS: 92

; SOFTWARE: Patent In version 3.0

; SEQ ID NO 5

; LENGTH: 381

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-10-067-649-5

Query Match 21.4%; Score 479; DB 9; Length 381;

Best Local Similarity 29.9%; Pred No. 1.4e-35;

Matches 120; Conservative 78; Mismatches 144; Indels 60; Gaps 12;

QY 17 DHNLTRQFIALYRLRPLVYTPPELPG-----RAKIALVITGVLIIFALALFGN 63

Db 9 DHNLTRQFIALYRLRPLVYTPPELPG-----RAKIALVITGVLIIFALALFGN 68

QY 64 ALVYVYVTRSKAMRTVTNIFCSIALSDLLITFFCIPTVMTLQNSIDNMLGAGFICKMVPV 123

Db 64 ALVYVYVTRSKAMRTVTNIFCSIALSDLLITFFCIPTVMTLQNSIDNMLGAGFICKMVPV 123

QY 69 SLVHVIVIKFSMRTVTNFFIANLAVADLLVNTLCIPLPTFTYTLMGKMKWCPVLCHLVY 128

Db 69 SLVHVIVIKFSMRTVTNFFIANLAVADLLVNTLCIPLPTFTYTLMGKMKWCPVLCHLVY 128

QY 124 VQSTAVVTEILTMTCTIAVERHGLVHPEKMKQWYTNRAFTMLGVWLVAVIVGSPMHV 183

Db 124 VQSTAVVTEILTMTCTIAVERHGLVHPEKMKQWYTNRAFTMLGVWLVAVIVGSPMHV 186

QY 129 AGLAVQVSTITLTVIALCPHCVIYHLESK---ISKQISELIIIGLAWGVSALLASPLAIF 186

Db 129 AGLAVQVSTITLTVIALCPHCVIYHLESK---ISKQISELIIIGLAWGVSALLASPLAIF 186

QY 184 QO---LEIKYDPLVYEKEHICGLEW---TSPVHOKIYTTFTLIVTLFLLPMLVMLLYSKI 237

Db 184 QO---LEIKYDPLVYEKEHICGLEW---TSPVHOKIYTTFTLIVTLFLLPMLVMLLYSKI 237

QY 187 REYSLIEIPDF---EIVACTEKWPGEEKSVYGVTSLSLTLLIYVPLGLIISFSYTRI 242

Db 187 REYSLIEIPDF---EIVACTEKWPGEEKSVYGVTSLSLTLLIYVPLGLIISFSYTRI 242

QY 238 GYELW---TKRVGDSGLVLTTHGKEMSKIAKKKRAVIMVTVVAFVAVCWAPPHVHHMM 295

Db 238 GYELW---TKRVGDSGLVLTTHGKEMSKIAKKKRAVIMVTVVAFVAVCWAPPHVHHMM 295

QY 243 ---WSKIKNHVSPGAASDHYH-----QRRHKTKMLVCVWVFAVSWLPLHAPOLA 290

Db 243 ---WSKIKNHVSPGAASDHYH-----QRRHKTKMLVCVWVFAVSWLPLHAPOLA 290

QY 296 IEYSNFE---KEYDDVTIKMIFAIQVIGFSNSICNPVYAFNENEFKNVLSAVCYCIV 352

Db 296 IEYSNFE---KEYDDVTIKMIFAIQVIGFSNSICNPVYAFNENEFKNVLSAVCYCIV 352

QY 291 VQIDSHVLDKEY-----KLFTVFHIIAMGSTANFLLYGMWNSNTRKAFLSA----- 339

Db 291 VQIDSHVLDKEY-----KLFTVFHIIAMGSTANFLLYGMWNSNTRKAFLSA----- 339

QY 353 NKTFSPAQR---HCNSGITMMRKKAKSLRENVEETKGA 390

Db 353 NKTFSPAQR---HCNSGITMMRKKAKSLRENVEETKGA 390

QY 340 ---PRCEQRLDAIHSEVSMTEFKAKKNLEVKKNKGLIDSFSEA 378

Db 340 ---PRCEQRLDAIHSEVSMTEFKAKKNLEVKKNKGLIDSFSEA 378

# RESULT 8

US-09-962-64b-1b

; Sequence 16, Application US/09962646

; Patent No. US20020103123A1

; GENERAL INFORMATION:

; APPLICANT: GERALD, CHRISTOPHE P.G.

; APPLICANT: WEINSHANK, RICHARD I.

; APPLICANT: WALKER, MARY W

; APPLICANT: BRANCHER, THERESA

; TITLE OF INVENTION: MODIFYING FEEDING BEHAVIOR, COMPOUNDS USEFUL IN SUCH METHODS,

; FILE REFERENCE: 1795/46166BZA

; CURRENT APPLICATION NUMBER: US/09/962,646

; CURRENT FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: 09/200,673

; PRIOR FILING DATE: 1998-11-25

; PRIOR APPLICATION NUMBER: 08/566,096

; PRIOR FILING DATE: 1995-12-01

; PRIOR APPLICATION NUMBER: 08/349,025

; PRIOR FILING DATE: 1994-12-01

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: Patent In version 3.1

; SEQ ID NO 16

[illegible]

RESULT 11  
 US 09-992-331 18  
 ? Sequence 18, Application US/09992331  
 ? Publication No. US2003022186A1  
 ? GENERAL INFORMATION:  
 ? APPLICANT: FEDER, JOHN N.  
 ? APPLICANT: MINTIER, GABE  
 ? APPLICANT: RAMANATHAN, CHANDRA S.  
 ? APPLICANT: HAWKEN, DONALD R.  
 ? TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,  
 ? TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA  
 ? TITLE OF INVENTION: CELLS  
 ? FILE REFERENCE: D0048NP  
 ? CURRENT APPLICATION NUMBER: US/09/992,331  
 ? CURRENT FILING DATE: 2001-11-14  
 ? PRIOR APPLICATION NUMBER: 60/308,540  
 ? PRIOR FILING DATE: 2001-07-27  
 ? PRIOR APPLICATION NUMBER: 60/261,782  
 ? PRIOR FILING DATE: 2001-01-16  
 ? PRIOR APPLICATION NUMBER: 60/248,483  
 ? PRIOR FILING DATE: 2000-11-14  
 ? NUMBER OF SEQ ID NOS: 45  
 ? SOFTWARE: Patent In Ver. 2.1  
 ? SEQ ID NO 18  
 ? LENGTH: 385  
 ? TYPE: PRT  
 ? ORGANISM: Mus musculus  
 US-09-992-331-18

Query Match 20.9%, Score 466.5, DB 9, Length 385;  
 Best Local Similarity 30.6%, Pred. No. 2c-34,  
 Matches 117; Conservative 71; Mismatches 121; Indels 73; Gaps 13;  
 QY 38 PELPGRAKLA-----LVITGLVIFALALFGNALVYVYVTRSKAMRTVTNIFICSLSLSD 92  
 DB 42 PELIDSTKLVEQVQLHAYGSIITLLGVGNSLVHVVIFKFSKMTVTNFFIANLAVADI 101  
 QY 93 LITTFICIPVTMLONISDNWLGAPICKMVPFVOSTAVVTELLTMTCLAVRHOGI.VHPFK 152  
 DB 102 LVNTLCLEFTLYITLNGEKMGVPLCHLVPAQGLAVOVSTITLTVIALDRHRCIVVHLE 161  
 QY 153 KMWQYINRRAPITMLGVVVLAVVGVSPMWHVQV---LPIKYDFLYEKEHICCLPEW---T 206  
 DB 162 SK -LSKRISPLIGLAWGTSALLASPLAIFREYSLFIIPDF---EIVACTEKWIGEE 215  
 QY 207 SPVHQIKYITFTILVTLFLLPMLVILYLSKIGYELWIKKR--VGDGSLVRLTIHGKEMSKI 264  
 DB 216 KSVYGVTVYSLTLLLYVLPGLITTSFSYTRI---WSKLRNHVSPGAASDHYH----- 264  
 QY 265 ARKKKRAVIMVTVVALFAVWAPPHVVMHMEYSNFE--KEYDDVTIKMIFAVOIIIG 321  
 DB 265 QRRUKMTKMLVGVVVVFAVSWLPLHAFQLAVDIDSHVLDLKEY-----KLITVFHIIA 318  
 QY 322 FSNISICNPVIAFMHNEPKKVLASVYCYIVNKTFTSPAQR---HNSGUTIMRKKKAKFS 377  
 DB 319 MGSFTANPLLYGMMSNRYRKAFISA-----FRCQRLDATHSEKVSMTFRACK----- 365  
 QY 378 IRENPEETKGEAFSDKNIIEVK 399  
 DB 366 -----NLEVK 470

RESULT 12  
 US 09-992 331-19  
 ? Sequence 19, Application US/09992331  
 ? Publication No. US2003022186A1  
 ? GENERAL INFORMATION:  
 ? APPLICANT: FEDER, JOHN N.  
 ? APPLICANT: MINTIER, GABE  
 ? APPLICANT: RAMANATHAN, CHANDRA S.  
 ? APPLICANT: HAWKEN, DONALD R.  
 ? TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,  
 ? TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA

? TITLE OF INVENTION: CELLS  
 ? FILE REFERENCE: D0048NP  
 ? CURRENT APPLICATION NUMBER: US/09/992,331  
 ? CURRENT FILING DATE: 2001-11-14  
 ? PRIOR APPLICATION NUMBER: 60/308,540  
 ? PRIOR FILING DATE: 2001-07-27  
 ? PRIOR APPLICATION NUMBER: 60/261,782  
 ? PRIOR FILING DATE: 2001-01-16  
 ? PRIOR APPLICATION NUMBER: 60/248,483  
 ? PRIOR FILING DATE: 2000-11-14  
 ? NUMBER OF SEQ ID NOS: 45  
 ? SOFTWARE: Patent In Ver. 2.1  
 ? SEQ ID NO 19  
 ? LENGTH: 444  
 ? TYPE: PRT  
 ? ORGANISM: Homo sapiens  
 US-09-992-331-19

Query Match 20.4%, Score 455, DB 9, Length 444;  
 Best Local Similarity 27.7%, Pred. No. 2.7c-33;  
 Matches 104; Conservative 91; Mismatches 142; Indels 38; Gaps 9;  
 QY 4 LNIIEQESRLLRHNLITREQFIYALPLPLVYTPPELPFGAKLALVLTGVLIFALALFGN 63  
 DB 21 LNETQEPF---LNPITYDDDEEFLR-YLMREYLH----PKYEWVLIAGYIIVFVVALIGN 72  
 QY 64 ALVYVYVTRSKAMRTVTNIFICSLSLSDLLITFEICIPVTMLONISDNWLGAPICKMVPF 123  
 DB 73 VLVCVAVKKNHMTVTNFFIVNLSDVLVTITCLPATLVVDITETWTFGSLCKVIPY 132  
 QY 124 VUSTAVVTELLTMTCLAVRHOGI.VHPFKMKNWYTNRRAPITMLGVVVLAVVGVSPMWHV 183  
 DB 133 LQTVSVSVSVTLSCIALDRWYAICHPLMEK--STAKRARNISIVIIWVSCIIMIPQAIY 190  
 QY 184 QOLEIKYDFLYEKEHI---CCLEEMTSPVHQIKYITFTILVILELLPLMVLILYLSKIGYEL 241  
 DB 191 MKCSIVFVGLANKTLTFTVCDHRCWGGELVYKMYHICFFIYVWAPICLWVLATLQIFPKL 250  
 QY 242 WIKKVRQGSVLP-----TIHGKEMSKIARKKKRAVIMVTV 278  
 DB 251 WFRQTEGTSVVQVRKPKLPQVSPRGJPTKSPMSAVAAEIKQIRAPKPTAPMLMV-V 309  
 QY 279 VALFAVCWAPPHVVMHMEI-YSNFEKEYDDVTIKMIFAVOIIICFNSICNPVIAFMNE 337  
 DB 310 LLVFAICYLPISILNLVKRVEGMAFTEDETRETIVAWTFTHSLVLYANSANPIIYNFLSG 369  
 QY 338 NFKKNVLSAV-CYCI 351  
 DB 370 KFRREFKAAFSCCCL 384

RESULT 13  
 US-10-282-717-2  
 ? Sequence 2, Application US/10282717  
 ? Publication No. US20030083466A1  
 ? GENERAL INFORMATION:  
 ? APPLICANT: VANAGISAWA, MASASHI  
 ? TITLE OF INVENTION: CDNA CLONE MY1 THAT ENCODES A NOVEL  
 ? TITLE OF INVENTION: HUMAN 7-TRANSMEMBRANE RECEPTOR  
 ? FILE REFERENCE: GH50029D1C1  
 ? CURRENT APPLICATION NUMBER: US/10/282,717  
 ? CURRENT FILING DATE: 2002-10-28  
 ? PRIOR APPLICATION NUMBER: 09/676,625  
 ? PRIOR FILING DATE: 2000-10-02  
 ? PRIOR APPLICATION NUMBER: 09/119,788  
 ? PRIOR FILING DATE: 1998-07-21  
 ? PRIOR APPLICATION NUMBER: 60/053,790  
 ? PRIOR FILING DATE: 1997-07-25  
 ? NUMBER OF SEQ ID NOS: 2  
 ? SOFTWARE: FastSeq for Windows Version 3.0  
 ? SEQ ID NO 2  
 ? LENGTH: 444  
 ? TYPE: PRT

ORGANISM: Homo sapiens  
US-10-282-717-2

Query Match 20.3%; Score 453; DB 9; Length 444;

Best Local Similarity 27.7%; Pred No. 4 1e-33;

Matches 104; Conservative 90; Mismatches 143; Indels 38; Gaps 9;

QY 4 LNIETPEQSKLLRDNHNTREOFIALYRLRPLVYTPPELGRKALALVLTGVLIFALFALFQ 63  
DB 21 LNETQEPF---LNTDVEDDEFLR-YLMREYLH-----PKYEYEWLIAGYIIVFVALIGN 72  
QY 64 ALVYVVYVTRSKAMRTVNTIFCSIALSDLLITFFCPVIMLONIISNMLGGAFICKMVPF 123  
DB 73 VLVCVAVYKNNHMTVNTNYFINLSLADVLVTITCLPATLVVDITETWFGQSICKVDPY 132  
QY 124 VQSTAVYTEILTMTCIAVERHOGVLVHPFKMKQVYTNRAFTMLGVVWLVAIVVGSPPWHV 183  
DB 133 LQTVSVSVSVLTSLADRWYAICHPLMFK--STAKRARNISVITWIVSCIIIMIPQAI 190  
QY 184 QOLEIKYDFLYEKKEHI--CCLEEWTSVPHOKIYTFILVILFLLPLMVMILYLSKIGYEL 241  
DB 191 MECSTVPFGLANKTTLTVTDERWCGEITYPKMWHICEFLVITYMAPLCMLVLAQLPRKL 250  
QY 242 WIKKRVGDSGLRT-----JHGKEMSKIARKKKRAVIMMVTV 278  
DB 251 WCRQIPGTSSVYVOKKKWKLQHPVSQPRGQGP-IKSRMGAVAAEIKQIKARKKKTAKMLMV-V 309  
QY 279 VALFAVCWAPHHVYVHMIE-YSNPEKEYDDVTIKMIFAIVQIIGFSNSICNPVIYAFMNE 337  
DB 310 LLVTAICYLPIISILNVLKRVGMPAHTDRETVYAMFTFSHMLVYANSAANPIIYNPLSG 369  
QY 338 NFKKNVLSAV-CYCI 351  
DB 370 KFEDEFKAESCCCL 384

# RESULT 14

US-10-067-649-6

Sequence 6, Application US/10067649

Publication No. US20030100057A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company

TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY14, RELATED TO

FILE REFERENCE: D0118 NP

CURRENT APPLICATION NUMBER: US/10/067.649

CURRENT FILING DATE: 2002-03-05

PRIOR APPLICATION NUMBER: US 60/266,525

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: US 60/329,897

PRIOR FILING DATE: 2001-10-16

NUMBER OF SEQ ID NOS: 92

SOFTWARE: PatentIn version 3.0

SEQ ID NO 6

LENGTH: 385

TYPE: PRT

ORGANISM: Gallus gallus

US-10-067-649-6

Query Match

Best Local Similarity 29.5%; Pred. No. 5.2e-32;

Matches 119; Conservative 81; Mismatches 129; Indels 74; Caps 16;

QY 23 EGFALYRLRLVLT-----PELGRKALALVLTGLVIFA-----LALFQNAL 65

DB 19 ELFTKLYLPR---VTPVSELALDPKPELKDSTTLVEVQI-LIIFAYCSIIILGVICNSL 74

QY 66 VYVYVTRSKAMRTVNTIFCSIALSDLLITFFCPVIMLONIISNMLGGAFICKMVPFVQ 125

DB 75 VHIIVIKSKMRTVNTIFANLAVADELLTCLPLFTVLTLLGKWLGPVLCHLPYQAQ 134

QY 126 STAVVTEILTMTCIAVERHOGVLVHPFKMKQVYTNRAFTMLGVVWLVAIVVGSPPWHVQ 185

DB 135 ALAVHVSIVTLTVIALDRHRCIVVYHLESK--ISKRIISFLIIGVAMAVSALLASPLAIPRE 192  
QY 186 ---LEIKYDFLYEKKEHICCLEEWTSPP---VHQKIYTFILVILFLLPLMVMILYLSKIGY 239  
DB 193 YSLIEIIPDF---KIVVCEKWPGEQLNGYIYSVSMIIQYVLPFLAIIYSAYTRI-- 246  
QY 240 ELW--IKKRVGDSGLRTIHKEMSKIARKKKRAVIMMVTVVAVLFAVCWAPHHVYVHMIE 297  
DB 247 --WTKLNHVSPPGAGNHHYH-----HKKQITTKMLVCVVVVVFAVSWLFPFAHCLVSD 296  
QY 298 YSNFE---KEYDDVTIKMIFAIVQIIGFSNSICNPVIYAFMNEFKKNVLSAVCYCIVNK 354  
DB 297 IDSQVLLDKY-----KLIVTFEHVIAMCSTFANPLLYCMNNNRYTAFLLTA----- 343  
QY 355 TFSFAQRHGN-----SGITMMRKKAFLSLENPEETKEAFS 392  
DB 344 -FQEQRLDSTHPEVSAAFKARKKLEAKKSQFP-----GDSET 380

# RESULT 15

US-09-866-248A-2

Sequence 2, Application US/09866248A

Publication No. US20020198367A1

GENERAL INFORMATION:

APPLICANT: Gerald, Christophe P.G.

APPLICANT: Jones, Kenneth A.

APPLICANT: Bonini, James A.

APPLICANT: Borowsky, Beth

TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors

TITLE OF INVENTION: and Uses Thereof

FILE REFERENCE: 1795/57155-A

CURRENT APPLICATION NUMBER: US/09/866,248A

CURRENT FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: 09/161,113

PRIOR FILING DATE: 1998-09-25

NUMBER OF SEQ ID NOS: 42

SOFTWARE: Patent In Ver. 2.0 - beta

SEQ ID NO 2

LENGTH: 432

TYPE: PRT

ORGANISM: Rattus norvegicus

US-09-866-248A-2

Query Match 19.6%; Score 437; DB 9; Length 432.

Best Local Similarity 28.1%; Pred. No. 1.1e-31;

Matches 101; Conservative 70; Mismatches 119; Indels 70; Gaps 8.

QY 11 FSRILLRDNHNTREOFIALYRLRPLVYTPPELGRKALALVLTGVLIFALFALFALVYV 70

DB 32 FSSYYOHSSPVAAFMFAAY-----VLIIFLCMVGNITLVCFIV 68

QY 71 TRSKAMRTVNTIFCSIALSDLLITFFCPVIMLONIISNMLGGAFICKMVPFVSTAVV 140

DB 69 LKNRHMRTVNTMFIINLAVSDLLVIGPCMTTIVDNLITGCPDNATCKMGLVOGMSVS 128

QY 131 TEILTMTCIAVERHOGVLVHPFKMKQVYTNRAFTMLGVVWLVAIVVGSPPWHVQOOLEIKY 190

DB 129 ASVETLVAIAVERFCIVHPPREK--UTLRKALFTTAVIWAALLIMCPS-----AVTL 180

QY 191 DFLEYEKKEHI-----CCLEEWTSVPHOKIYTFILVILFLLPLMVMILYLSKIG 238

DB 181 TVTREHHFMDLDAKNSRPIYSCWEAPKGMKKVYTAIVLFAHYILVPLAIIIVMYVYRIA 240

QY 230 YELW-----IFKPVWISVLPITIHGKEMSKIARKKKRAVIMMVTVVAVLFAVWAPFH 290

DB 241 PKICQAPCPARTHEAFAEG-----GPTSPRRARVVHMLVMVALFTLISLPIW 289

QY 291 VYHMMIYSNPEKEYDDVTIKMI-----FAIVQIIGFSNSICNPVIYAFMNEFKKNVISA 346

DB 290 VILLIIDYG-----ELSELQHLHLSVYAFPLAHLAFTFSSANPIIYGYNENFERGQAA 345

Search completed: June 10, 2003, 13:12:18

Job time : 27 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2003, 06:07:46 ; Search time 3547 seconds  
(without alignments)  
3496.880 Million cell updates/sec

Title: US-10-070-241b-1

Perfect score: 2235

Sequence: 1 MQALNITPEQFSRLLRHNL PHILFRSELAENSPDLSGH 431

Scoring table: RIQSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4100280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n\_model -DRV=xlh  
-O=/cgn2.1/USPTO.spool/US10070241/runat\_10062003\_095453\_20838/3pp\_query\_fasta\_1.583  
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOWE=L=0 -LOWEEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 codi -LIST=45  
-DOCALIGN=200 -THP\_SCORE=pt -THP\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-USER=US10070241 -ACGN\_1\_1\_2496 -runat\_10062003\_095453\_20838 -NCPU=6 -CPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -NSPLOCK=100 -LONGLOG  
-DEV\_TIME=120 WAPN\_TIME=30 THPEADS=1 XGAPCP=10 XGAPEXT=0.5 XGAPCP-6 -XGAPEXT-7  
-XGAPEXT=7 -YGAPCP=10 -YGAPEXT=0.5 -DELCP=6 -DELEXT=7

Database :

GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_em.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_mus.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*

- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rnd.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description               |
|------------|-------|---------------|--------|-------|---------------------------|
| 1          | 2235  | 100.0         | 1293   | 6     | BD012995 Novel G p        |
| 2          | 2235  | 100.0         | 1293   | 23    | BD010028 Novel G p        |
| 3          | 2235  | 100.0         | 1296   | 6     | AX360195 Sequence         |
| 4          | 2235  | 100.0         | 1713   | 6     | AX360197 Sequence         |
| 5          | 1972  | 88.2          | 1368   | 9     | AF411117 Homo sapi        |
| 6          | 741.5 | 33.2          | 155347 | 2     | AC005961 Homo sapi        |
| c 7        | 735   | 32.9          | 75164  | 9     | AC095867 Homo sapi        |
| 8          | 711   | 31.8          | 432    | 6     | AX277359 Sequence         |
| c 9        | 711   | 31.8          | 432    | 6     | AX277359 Sequence         |
| c 10       | 711   | 31.8          | 432    | 6     | AX277359 Sequence         |
| c 11       | 579   | 25.9          | 170227 | 9     | AC093816 Homo sapi        |
| 12         | 508   | 22.7          | 1264   | 10    | AF268500 Rattus no        |
| 13         | 502   | 22.5          | 1263   | 9     | AF268899 Homo sapi        |
| 14         | 502   | 22.5          | 1263   | 9     | AF330053 Homo sapi        |
| 15         | 502   | 22.5          | 1560   | 9     | AF257210 Homo sapi        |
| 16         | 502   | 22.5          | 1888   | 6     | E32646 7:M. r. c. p. t. o |
| 17         | 501.5 | 22.4          | 1936   | 9     | AF119815 Homo sapi        |
| 18         | 496   | 22.2          | 1870   | 9     | AF236083 Homo sapi        |
| 19         | 491.5 | 22.1          | 1412   | 5     | AF000245 Homo sapi        |
| 20         | 487   | 21.8          | 1254   | 10    | AF330054 Mus muscu        |
| 21         | 486   | 21.7          | 1556   | 6     | AF000041 Sequence         |
| 22         | 485.5 | 21.7          | 1315   | 5     | AF073925 Cadus mer        |
| 23         | 479.5 | 21.5          | 1286   | 5     | AF037400 Danio rer        |
| 24         | 479   | 21.4          | 1147   | 10    | AY004257 Rattus no        |
| 25         | 477.5 | 21.4          | 1228   | 10    | AF236084 Mus muscu        |
| 26         | 475   | 21.3          | 129767 | 2     | AC109773 Rattus no        |
| 27         | 474   | 21.2          | 1203   | 6     | AF388042 Sequence         |
| 28         | 473.5 | 21.2          | 3379   | 10    | D86238 Mus musculu        |
| c 29       | 473.5 | 21.2          | 214857 | 10    | AC122899 Mus muscu        |
| c 30       | 473.5 | 21.2          | 254775 | 2     | AC091478 Mus muscu        |
| 31         | 469.5 | 21.0          | 1200   | 6     | AX138820 Sequence         |
| 32         | 469.5 | 21.0          | 1200   | 6     | AX139005 Sequence         |
| 33         | 469.5 | 21.0          | 1200   | 6     | AX139424 Sequence         |
| 34         | 469.5 | 21.0          | 1200   | 6     | AX463056 Sequence         |
| 35         | 469.5 | 21.0          | 1200   | 9     | HS036269 Human neuro      |
| 36         | 469.5 | 21.0          | 1200   | 6     | AF000040 Sequence         |
| 37         | 469.5 | 21.0          | 1200   | 6     | 124737 Sequence 1         |
| 38         | 469.5 | 21.0          | 1479   | 4     | BT050144 Bos taurus       |
| 39         | 469.5 | 21.0          | 3444   | 9     | HSNRY2S2 Human type       |
| 40         | 469.5 | 21.0          | 3747   | 9     | HS042766 Human neuro      |
| 41         | 469.5 | 21.0          | 102822 | 9     | AC104407 Homo sapi        |
| 42         | 468.5 | 21.0          | 1134   | 5     | AF410853 Gallus ga        |
| 43         | 468.5 | 21.0          | 1146   | 9     | AF303090 Musaca mu        |
| 44         | 468.5 | 21.0          | 1240   | 9     | HS076254 Human neuro      |
| 45         | 468.5 | 21.0          | 2144   | 6     | AF199657 Sequence         |

ALIGNMENTS

RESULT 1

B0012995  
 LOCUS 1293 bp DNA linear PAT 02 AUG-2002  
 DEFINITION Novel G protein coupled receptor protein and its DNA.  
 ACCESSION B0012995  
 VERSION B0012995.1 GI:22093184  
 KEYWORDS NO 0116316.A/1  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata, Vertebrata, Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini, Hominoidea; Homo.  
 Watanabe, T., Kikuchi, K. and Shintani, Y.  
 REFERENCES Novel G protein coupled receptor protein and its DNA  
 PATENT: WO 0116316-A 1 08-MAR-2001.  
 JOURNAL TAKEHIA CHEMICAL INDUSTRIES LTD, TAKUYA WATANABE, KUNIKO KIKUCHI,  
 YASUSHI SHINTANI  
 OS Homo sapiens (human)  
 PN WO 0116316 A/1  
 PD 08-MAR-2001  
 PF 24-AUG-2000 WO 2000P005684  
 PR 27-AUG-1999 JP 999 241530  
 PI TAKUYA WATANABE, KUNIKO KIKUCHI, YASUSHI SHINTANI PC  
 C12N15/12, C12N15/09, C07K14/705, C07K16/28, C12P21/02, A61K45/00, PC  
 A61P43/00.  
 CC G01N33/566, G01N33/15, G01N33/50  
 FH KEY Location/Qualifiers.  
 FEATURES Location/Qualifiers  
 1..1293  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 442 a 405 c 401 g 345 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4, 48e-196 Length: 1293  
 Score: 2235.00 Matches: 431  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US 10 070 241b 1 (1 431) x B0012995 (1 1293)  
 QY 1 MetGlnAlaLeuAsnIleThrProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20  
 DB 1 ATGGAGAGCGCTTAACATTACCGCGAGCAGTTCCTCCGGTGGTGGGAGCAAAACACCTG 50  
 QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40  
 DB 61 AGCGGGAGCAGCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 41 ProGlyArgAlaIleLysLeuAlaLeuValIleuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
 DB 121 GGGGAGCGAGCTAAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180  
 QY 61 PheGlyAsnAlaLeuValIlePheTyrValValIleThrArgSerLysAlaMetArgThrValThr 80  
 DB 181 TTTGGCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100  
 DB 241 AACATCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMet 120  
 DB 301 GTCAACATGCTCCAGACATCTCCGACAACTGGATGAGGAGTGGTGGTGGTGGTGGTGGTGG 360  
 QY 121 ValProPheValGlnSerThrAlaValIleThrGlnIleLeuThrMetThrCysIleAla 140  
 DB 361 GCGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 QY 141 ValGluAlaGlnIleGlyLeuValHisPropHelysMetLysTrpGlnTyrThrAsnArg 160

421 GTGAAAGGAG 480  
 QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
 DB 481 AGAGATTTTCAATGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 DB 541 TGGCAAGTGGCAACACTGACATCAAAATGACATTCCTTATATGAAAAGGAAACACATCTCC 600  
 QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleTyrThrThrPheIleLeuVal 220  
 DB 601 TGCCTTAGAAGAGTGGACAGGCTGTGGACCAAGAAATCATACACACTTCATCTGCTGCTGCTG 660  
 QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGlu 240  
 DB 661 ATCCCT 720  
 QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260  
 DB 721 CTTTGCATAAAGAAAGAGTGGGATGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCA 780  
 QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValValAla 280  
 DB 781 ATGTCACAAATAG 840  
 QY 281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsn 300  
 DB 841 CTCTTTTCT 900  
 QY 301 PheGluLysGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320  
 DB 901 TTGGAAAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
 QY 321 GlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLys 340  
 DB 961 GGATTTTCCAACTCCATCTGTAATCCCATTTCTATCTATTATGAATAAAAATTCAAA 1020  
 QY 341 CysAsnValLeuSerAlaValCysTyrCysIleValAsnLysIlePheSerProAlaGln 360  
 DB 1021 AAAAATGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
 QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380  
 DB 1081 AGGCATCGAAATTCAGGAATTCAGATCGCGAAGAAAGCAAGTTTCCCTCCACAGAG 1140  
 QY 381 AsnProValGluGluThrLysGlyGlyAlaPheSerAspGlyAsnIleGluValLysLeu 400  
 DB 1141 AATCGACTGGAGGAATTCAAAAGAGATTTATGATGATGATGATGATGATGATGATGATG 1200  
 QY 401 CysGluGlnThrGluGluLysLysLysLysArgHisIleuAlaLeuPheArgSerGlu 420  
 DB 1201 TGTGAGCAACAAG 1260  
 QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
 DB 1261 CTGGTTCAGAAATCTCTCTTTAGACAGTGGCAT 1293  
 RESULT 2  
 B0010028 standard; DNA; HUM; 1293 BP.  
 ID B0010028  
 XX AC HD010028;  
 XX SV BD010028.1  
 XX DT 08-FEB-2002 (Rel. 70, Created)  
 XX DT 08-FEB-2002 (Rel. 70, Last updated, Version 1)  
 XX DE Novel G protein coupled receptor protein and its DNA.  
 XX KW JP 03075483-T/1.





TITLE Human neuropeptide Y-like q protein-coupled receptor  
JOURNAL Patent: WO 0204518-A 17 JAN 2002;  
Bayer Corporation (US)

FEATURES  
source Location/Qualifiers  
1..1296

BASE COUNT 444 a 305 c 301 g 346 t  
ORIGIN /db\_xref="taxon:9606"

Alignment Scores:  
Pred. No.: 4,396-196 Length: 1296  
Score: 2235.00 Matches: 431  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-070-241b-1 (1-431) x AX360195 (1-1296)

QY 1 MetGluAlaLeuAsnIleThrProGluInPheSerArgLeuLeuArgAspHisAsnLeu 20  
DB 1 ATGCAGGGCTTAAATTAAATGAGAGAGATGCTCTGAGTGTGAGGAGAAATG 50  
QY 21 ThrArgGluInPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40  
DB 61 AGCGGGGAGAGATTCATGCTGTCTTACGCTGGTGGAGAGAGAGAGAGAGAGAG 120  
QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
DB 121 CAGGACGGCGCAAGTGGCGCTGCTGTCACCGCGGCGGCGGCGGCGGCGGCGGCTC 180  
QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80  
DB 181 TTGGCAATGCTGCTGCTTCTACGCTGTGAGATGAGAGAGAGAGAGAGAGAGAGAG 240  
QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPheCysIlePro 100  
DB 241 AACATCTTTATCTGCTCTGAGCTCAGTACCTGCTATCAGCTTCTTGGATTGCC 300  
QY 101 ValThrMetLeuGluAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMet 120  
DB 301 GTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 121 ValProPheValGluSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
DB 361 GTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160  
DB 421 GTGCAAG 480  
QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
DB 481 AGCGCTTTCACAAATGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 181 TrpHisValGlnIleLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
DB 541 TGCGAGTGCACCACTTGAATGATCAATATGACTTCTATATGAAAGAGACATCTGC 600  
QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleCysIleThrPheIleLeuVal 220  
DB 601 TGCTTAGAAGAGTGCAG 660  
QY 221 IleLeuPheLeuLeuProLeuMetValIleLeuIleLeuTyrSerLysIleGlyTyrGlu 240  
DB 661 ATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260  
DB 721 CTTTGATTAAG 780  
QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValAlaVal 280

DB 781 ATCTCCAAAATAGCCAGCAAGCAAAAATATGATATGATGAGAGAGAGAGAGAGAGAG 840  
QY 281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsn 300  
DB 841 CTCCTTTGCTGTGCTGGCCACCATTCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 301 PheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320  
DB 901 TTTGAAAGAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
QY 321 GlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLys 340  
DB 961 GGATTTTCCAACTCCATCTGTAATCCCATCTGCTATGCTATGCTATGCTATGCTATGCT 1020  
QY 341 LysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGln 360  
DB 1021 AAAAAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380  
DB 1081 AGGATGCAAAATTCAGCAATTAAATATATATATATATATATATATATATATATATATAT 1140  
QY 381 AsnProValGluGluThrLysGlyCysAlaPheSerAspGlyAsnIleGluValLysLeu 400  
DB 1141 AATCCAGTGGAG 1200  
QY 401 CysGluGlnThrGluGluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGlu 420  
DB 1201 TGTGAACACATAG 1260  
QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
DB 1261 CTGGCTAAATTTCTTTTAAATATATATATATATATATATATATATATATATATAT 1293  
RESULT 4  
AX360197  
LOCUS AX360197 1710 bp DNA linear PAT 13-FEB-2002  
DEFINITION Sequence 3 from Patent WO0204518.  
ACCESSION AX360197  
VERSION AX360197.1 GI:18675761  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
AUTHORS Bloomquist, B.T. and Zhelein, L.  
TITLE Human neuropeptide Y-like q protein-coupled receptor  
JOURNAL Patent: WO 0204518-A 3 17 JAN 2002;  
Bayer Corporation (US)  
FEATURES  
source Location/Qualifiers  
1..1710  
/db\_xref="taxon:9606"  
BASE COUNT 419 a 439 c 429 g 423 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,046-196 Length: 1710  
Score: 2235.00 Matches: 431  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-070-241b-1 (1-431) x AX360197 (1-1710)  
QY 1 MetGluAlaLeuAsnIleThrProGluInPheSerArgLeuLeuArgAspHisAsnLeu 20  
DB 1 ATGCAGGGCTTAAATTAAATGAGAGAGATGCTCTGAGTGTGAGGAGAAATG 50  
QY 21 ThrArgGluInPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40  
DB 61 AGCGGGGAGAGATTCATGCTGTCTTACGCTGGTGGAGAGAGAGAGAGAGAGAGAGAG 120  
QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
DB 121 CAGGACGGCGCAAGTGGCGCTGCTGTCACCGCGGCGGCGGCGGCGGCGGCGGCTC 180  
QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80  
DB 181 TTGGCAATGCTGCTGCTTCTACGCTGTGAGATGAGAGAGAGAGAGAGAGAGAGAG 240  
QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPheCysIlePro 100  
DB 241 AACATCTTTATCTGCTCTGAGCTCAGTACCTGCTATCAGCTTCTTGGATTGCC 300  
QY 101 ValThrMetLeuGluAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMet 120  
DB 301 GTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 121 ValProPheValGluSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
DB 361 GTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160  
DB 421 GTGCAAG 480  
QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
DB 481 AGCGCTTTCACAAATGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 181 TrpHisValGlnIleLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
DB 541 TGCGAGTGCACCACTTGAATGATCAATATGACTTCTATATGAAAGAGACATCTGC 600  
QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleCysIleThrPheIleLeuVal 220  
DB 601 TGCTTAGAAGAGTGCAG 660  
QY 221 IleLeuPheLeuLeuProLeuMetValIleLeuIleLeuTyrSerLysIleGlyTyrGlu 240  
DB 661 ATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260  
DB 721 CTTTGATTAAG 780  
QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValAlaVal 280





|               |                      |                           |               |                             |
|---------------|----------------------|---------------------------|---------------|-----------------------------|
| repeat_region | /rpt_family="L1"     | complement(1394, .1436)   | repeat_region | complement(59879, .70176)   |
| repeat_region | /rpt_family="Alu"    | 8670, 9016                | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="L1"     | complement(11114, .12988) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="L1"     | complement(14973, .15269) | repeat_region | /rpt_family="THE1"          |
| repeat_region | /rpt_family="Alu"    | complement(15449, .16027) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="L1"     | complement(17220, .17578) | repeat_region | /rpt_family="MER21"         |
| repeat_region | /rpt_family="MER1"   | 20083, .20381             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | complement(20519, .21168) | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="L1"     | complement(23102, .23498) | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="MLT2B2" | 23880, .24560             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="L1"     | 26231, .26519             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | 26636, .26920             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | 27478, .27760             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | 28463, .28745             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | 28923, .29159             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | 29659, .29774             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="L1"     | 30064, .30219             | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="L1"     | 31258, .31368             | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="MIR"    | complement(33992, .40282) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="Alu"    | complement(40832, .40942) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="MIR"    | 42058, .42689             | repeat_region | /rpt_family="MER25"         |
| repeat_region | /rpt_family="L1"     | 42735, .42794             | repeat_region | /rpt_family="MER25"         |
| repeat_region | /rpt_family="L1"     | complement(46638, .47151) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="MLT2B2" | 47293, .47351             | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="MER46"  | complement(48972, .49022) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="Alu"    | complement(49129, .49432) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="Alu"    | 50485, .50683             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="MER30"  | complement(51366, .51451) | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="L1"     | 57160, .57423             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | 57619, .57917             | repeat_region | /rpt_family="Alu"           |
| repeat_region | /rpt_family="Alu"    | complement(57986, .58102) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="Alu"    | complement(59941, .60025) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="HY3"    | 60057, .60352             | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="Alu"    | 63695, .63963             | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="Alu"    | 67457, .67895             | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="MER4"   | complement(69113, .69482) | repeat_region | /rpt_family="L1"            |
| repeat_region | /rpt_family="THE1"   |                           | repeat_region | /rpt_family="Alu"           |
|               |                      |                           | repeat_region | complement(140097, .140362) |

Alignment Scores:  
 Pred. No.: 1.72e-56  
 Score: 741.50  
 Percent Similarity: 20.02%  
 Best Local Similarity: 19.54%

Length: 155347  
 Matches: 249  
 Conservative: 6  
 Mismatches: 5





University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 5 (bases 1 to 75164)  
 Watson, R.  
 Direct Submission  
 Submitted (01 MAR 2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Sep 28, 2001 this sequence version replaced q1:15778786.  
 Genomic Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/qsc>  
 Contact: sapiens@genome.wustl.edu  
 Summary Statistics  
 Center project name: HNH0587P06  
 Drafting Center: WMRK

NOTICE: This sequence may not represent the entire insert of this  
 clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >=  
 40); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.  
 McPherson, Department of Genetics, Washington University, St. Louis  
 Mo. For additional information about the map position of this  
 sequence, see <http://genome.wustl.edu/qsc>

#### SOURCE INFORMATION:

The RP11 human HAC library was made from the blood of one male  
 donor, as described by Osoegawa, K., Woon, P.Y., Zhao, H., Frangen, E.,  
 Taten, M., Catanese, J.J., and de Jong, P.J. (1998) An improved  
 approach for construction of bacterial artificial chromosome  
 libraries. Genomics 51:1-8. The clone may be obtained either from  
 Research Genetics, Inc. (<http://www.resgen.com>) or Peter de Jong  
 and coworkers at <http://www.chori.org>  
 VECTOR: pBAC3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-364P2, 2000 bp overlap.  
 Actual start of this clone is at base position 1 of RP11-587P6.

Polymorphisms have been identified between AC093867 and AC093816.

The sequence of AC088256 has been incorporated into AC093867

#### FEATURES

##### source

Location/Qualifiers  
 1..75164  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 /map="4"

/clone="RP11-587P6"  
 /clone\_lib="RP11-11"  
 995..1208

##### repeat\_region

/rpt\_family="MIR"

##### repeat\_region

1996..2018  
 /rpt\_family="AT-rich"

##### repeat\_region

2000..2238  
 /rpt\_family="L1"

##### repeat\_region

2239..2279  
 /rpt\_family="L1"

##### repeat\_region

4734..4872  
 /rpt\_family="ERV1"

##### repeat\_region

5098..5417

repeat\_region  
 5293..5315  
 /rpt\_family="Alu"

repeat\_region  
 5381..5419  
 /rpt\_family="AT-rich"

repeat\_region  
 5781..5859  
 /rpt\_family="L2"

repeat\_region  
 5875..5900  
 /rpt\_family="AT-rich"

repeat\_region  
 6593..8628  
 /rpt\_family="L1"

repeat\_region  
 14723..14746  
 /rpt\_family="AT-rich"

repeat\_region  
 14780..16396  
 /rpt\_family="L1"

repeat\_region  
 16197..16695  
 /rpt\_family="Alu"

repeat\_region  
 16696..16811  
 /rpt\_family="L1"

repeat\_region  
 16812..17096  
 /rpt\_family="Alu"

repeat\_region  
 17097..17643  
 /rpt\_family="L1"

repeat\_region  
 17644..17950  
 /rpt\_family="Alu"

repeat\_region  
 17951..18627  
 /rpt\_family="L1"

repeat\_region  
 18628..18936  
 /rpt\_family="Alu"

repeat\_region  
 18909..18938  
 /rpt\_family="L1"

repeat\_region  
 18937..19098  
 /rpt\_family="L1"

repeat\_region  
 19099..19347  
 /rpt\_family="Alu"

repeat\_region  
 19318..19351  
 /rpt\_family="L1"

repeat\_region  
 19348..21143  
 /rpt\_family="L1"

repeat\_region  
 20948..20984  
 /rpt\_family="AT-rich"

repeat\_region  
 21093..21123  
 /rpt\_family="AT-rich"

repeat\_region  
 21399..21572  
 /rpt\_family="MIR"

repeat\_region  
 23357..23448  
 /rpt\_family="MIR"

repeat\_region  
 24060..24115  
 /rpt\_family="L1"

repeat\_region  
 25558..25587  
 /rpt\_family="L1"

repeat\_region  
 25985..26006  
 /rpt\_family="AT-rich"

repeat\_region  
 26666..26822  
 /rpt\_family="L1"

repeat\_region  
 26869..27277  
 /rpt\_family="ERV1"

repeat\_region  
 27274..27355  
 /rpt\_family="L1"

repeat\_region  
 27393..27729  
 /rpt\_family="L1"

repeat\_region  
 27844..28060  
 /rpt\_family="L1"

repeat\_region  
 28178..28419  
 /rpt\_family="L1"

repeat\_region  
 28501..28599  
 /rpt\_family="L1"

repeat\_region  
 29208..30002  
 /rpt\_family="L1"

repeat\_region  
 30003..30100  
 /rpt\_family="Alu"

repeat\_region  
 30301..30544  
 /rpt\_family="L1"





```
Db 55333 GSCCGCTTAAATCATGTTTTAAAGATATATCAGGCGCGGTGCTGGCTCACACCTGGAA 55274
QY 265 ----- 265
Db 55273 TCCAGCACTTTGGGAGGCTTAAGTTGGAAATTTGTTGAGGCGAGAGTTTGAACGAG 55214
QY 265 ----- 265
Db 55213 CTTGGGCAACAAGTGAGACGCGCTCTCTACAAAAAATTTAAAAATTAGCTGGGTAGT 55154
QY 265 ----- 265
Db 55153 GRTATGACCTCTGGTCTGAGCTATGATGACAGGCTGAGGCTGGAGGATCACTTATGAGCC 55094
QY 265 ----- 265
Db 55093 CAGTTGAGGCTTCAGTGAGGATGATTGCAACACCTCTACACCTGAGGCAAGGAGT 55034
QY 265 ----- 265
Db 55033 AAGAGCTTGTCTCAAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 54974
QY 265 ----- 265
Db 54973 TATGATTTAAATTAATCCGCAACTTCCATATTTTCATTAATTTTTTTTGGTAGAGA 54914
QY 265 ----- 265
Db 54913 TGGTTTTGATGCTGCTAGGCTGCTGTTGAACCTCTGGAATCACTAAATTCCTCG 54854
QY 265 ----- 265
Db 54853 CCGTGTCTGCTCAAGAGCTGGATTAAAGSGTGAATCACTGCACTGGGCCATATTTT 54794
QY 265 ----- 265
Db 54793 ATAAATATGATTTTACTACACAGCATCAATGCTCTTTAATATAAAGTGACTAAGA 54734
QY 265 ----- 265
Db 54733 GTTTCATATTGAAGTCAAGAGTTAGACTTCTAATTCATCCCAAGATACAGCTTT 54674
QY 265 ----- 265
Db 54673 ACTAGCATCTAATTTGCAAAATTAATTAAGCTTCTGTGCTCTCAATTTCTGCATCTAT 54614
QY 265 ----- 265
Db 54613 CAAARCGCTTATTTGNTAAATTAATGACATTCATCTCATGCTATGACAGTGCCA 54554
QY 265 ----- 265
Db 54553 GATCATATATGAGTACAATTCAAGAACATTTATGATGTGAGGGTAATATCAATAGGAA 54494
QY 265 ----- 265
Db 54493 AAGATGATAATCAGGCAAGCAATACCTATCTGAGCTTAATTTAAAAAATATCTTGAG 54434
QY 265 ----- 265
Db 54433 TGAATTAAGCTCACACTTTAATATCTTTATCAAGCTGGCATTTGGCAAGGAGAAG 54374
QY 266 ----- 275
Db 54373 TTAAGTATAATTAAGTTCTGCTGGTTTAGGAAAGAGAAAGAGCTGCTATTATGATG 54314
QY 276 ValThrValAlaLeuPheAlaValICysIrrpAlaIlePheHisValAlaHisMetMet 295
Db 54313 GTGACAGTGGTGGCTCTCTTTGCTGCTGGCTGGGAACATTCTATGTTGCTCATATG 54254
QY 296 IleGluIysSer 299
Db 54253 ATTGAATACATGAGTGGTTGTCATTTCTCATTTGATAGACATTTCTCATTTCTCGCT 54194
```

```

RESULT 8
AX277359
LOCUS AX277359 432 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 17 from Patent WO0174897.
ACCESSION AX277359
VERSION AX277359.1 GI:16548924
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Vernet,C.A., Burgess,C.E., Fernandes,E., Taupier,P.T., Quinn,K.E.,
Szytek,K.A., Rastelli,L. and Herrmann,J.L.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0174897-A 17 11-OCT-2001;
Curagen Corporation (US)
FEATURES
source location/Qualifiers
1..432 /organism="Homo sapiens"
16..300 /db_xref="taxon:9606"
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10338.1"
/db_xref="GI:16548925"
/translation="MLGVVVLVAVIVGSPMHVQQLLEIKYDFLYEKEHICCLEWTSP
VHOKIYTFILVILFLMEBETSCHYDSDSGSLCCVLGTIPCCPYDD"
BASE COUNT 119 a 92 c 92 g 129 t
ORIGIN
Alignment Scores:
Pred. No. 1 24e-56 Length: 432
Score: 711.00 Matches: 143
Percent Similarity: 78.57% Conservative: 0
Best Local Similarity: 78.57% Mismatches: 0
Query Match: 31.81% Indels: 39
DB: 6 Gaps: 2
US-10-070-241B-1 (1-431) x AX277359 (1-432)
QY 160 ArgArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerPro 179
DB 1 CGAAGGCGTTTCACATGCTAGGTGGTGGCTGGCTGGCACTATCGTAGGATCACCC 60
QY 180 MetTrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIle 199
DB 61 ATGTGGCACCTGCAACAACITGAGATCAAAATAGACTTCTATATGAAAGGAACACATC 120
QY 200 CysCysLeuGluGluTrpThrSerProValHisGlnLysIleTyrThrThrPheIleLeu 219
DB 121 TGTGCTTAGAAGAGTGGACCGCTGTGCACGAGAAGATGTACACCACTTCATCCTT 180
QY 220 ValIleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyr 239
DB 181 GTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201
QY 240 GluLeu-TrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLy 259
DB 202 ---CTTATGG----- 208
QY 259 sGluMetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValVa 279
DB 209 -----AAGAAGAAACAGCTGTCATTATGATGGTACAGTGGT 246
QY 279 lAlaLeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSe 299
DB 247 GGCTCTCTTCTGCTGGCTGGCCACCATTCATGTTGTCATATGATGATGATGATGATGATG 306
QY 299 rAsnPheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGlnIle 319
DB 307 TAAATTTTGAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 366

```

---

```

QY 319 elieGlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPh 339
DB 367 TATGATGATTTTCAACTGCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 426
QY 339 eLys 340
DB 427 CAAA 430
RESULT 9
AX277395/c
LOCUS AX277395 432 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 53 from Patent WO0174897.
ACCESSION AX277395
VERSION AX277395.1 GI:16548960
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Vernet,C.A., Burgess,C.E., Fernandes,E., Taupier,P.T., Quinn,K.E.,
Szytek,K.A., Rastelli,L. and Herrmann,J.L.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0174897-A 53 11-OCT-2001;
Curagen Corporation (US)
FEATURES
source location/Qualifiers
1..432 /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 129 a 92 c 92 g 119 t
ORIGIN
Alignment Scores:
Pred. No. 1 24e-56 Length: 432
Score: 711.00 Matches: 143
Percent Similarity: 78.57% Conservative: 0
Best Local Similarity: 78.57% Mismatches: 0
Query Match: 31.81% Indels: 39
DB: 6 Gaps: 2
US-10-070-241B-1 (1-431) x AX277395 (1-432)
QY 160 ArgArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerPro 179
DB 432 CGAAGGCGTTTCACATGCTAGGTGGTGGCTGGCTGGCACTATCGTAGGATCACCC 373
QY 180 MetTrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIle 199
DB 372 ATGTGGCACCTGCAACAACITGAGATCAAAATAGACTTCTATATGAAAGGAACACATC 313
QY 200 CysCysLeuGluGluTrpThrSerProValHisGlnLysIleTyrThrThrPheIleLeu 219
DB 312 TGTGCTTAGAAGAGTGGACCGCTGTGCACGAGAAGATGTACACCACTTCATCCTT 253
QY 220 ValIleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyr 239
DB 252 GTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 232
QY 240 GluLeu-TrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLy 259
DB 231 ---CTTATGG----- 225
QY 259 sGluMetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValVa 279
DB 224 -----AAGAAGAAACAGCTGTCATTATGATGGTACAGTGGT 187
QY 279 lAlaLeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSe 299
DB 186 GGCTCTCTTCTGCTGGCTGGCCACCATTCATGTTGTCATATGATGATGATGATGATGATG 127
QY 299 rAsnPheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGlnIle 319
DB 126 TAAATTTTGAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 67

```



Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frenken, E., Tateno, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
VECTOR: pRACE3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-587P6. Actual start of this clone is at base position 73165 of RP11-587P6; actual end is at base position 170227 of RP11-364P2.

Data from AC093867 was used to finish this clone, AC093816.

Polymorphism has been identified between AC093867 and AC093816.

The sequence of AC055821 has been incorporated into AC093816.

#### FEATURES

##### source

Location/Qualifiers  
1..170227  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="4"

/map="4"  
/clone="RP11-364P2"  
/clone\_lib="RPCI-11"  
1..230

##### repeat\_region

/rpt\_family="L1"

##### repeat\_region

1644..1757  
/rpt\_family="L1"

##### repeat\_region

1762..1946  
/rpt\_family="L1"

##### repeat\_region

2511..2587  
/rpt\_family="L1"

##### repeat\_region

2803..3200  
/rpt\_family="MaLR"

##### repeat\_region

3216..3246  
/rpt\_family="AT\_rich"

##### repeat\_region

3249..6905  
/rpt\_family="L1"

##### repeat\_region

6416..6459  
/rpt\_family="AT\_rich"

##### repeat\_region

6907..6927  
/rpt\_family="AT\_rich"

##### repeat\_region

6920..7561  
/rpt\_family="L1"

##### repeat\_region

7562..7651  
/rpt\_family="MaLR"

##### repeat\_region

7704..7859  
/rpt\_family="MaLR"

##### repeat\_region

8069..8089  
/rpt\_family="AT\_rich"

##### repeat\_region

8260..8422  
/rpt\_family="MERL\_type"

##### repeat\_region

8903..9108  
/rpt\_family="MERL\_type"

##### repeat\_region

9164..9184  
/rpt\_family="AT\_rich"

##### repeat\_region

9484..9708  
/rpt\_family="MIR"

##### repeat\_region

9903..10288  
/rpt\_family="ERVL"

##### repeat\_region

10288..10471  
/rpt\_family="(TA)n"

##### repeat\_region

10310..10396  
/rpt\_family="ERVL"

##### repeat\_region

10474..10527

repeat\_region  
11244..11372  
/rpt\_family="MERL\_type"  
repeat\_region  
11373..11515  
/rpt\_family="L12"  
repeat\_region  
11581..11659  
/rpt\_family="L12"  
repeat\_region  
12197..12356  
/rpt\_family="MERL\_type"  
repeat\_region  
13088..13442  
/rpt\_family="L1"  
repeat\_region  
13706..14006  
/rpt\_family="Alu"  
repeat\_region  
15009..15322  
/rpt\_family="Alu"  
repeat\_region  
15294..15327  
/rpt\_family="(A)n"  
repeat\_region  
15877..16112  
/rpt\_family="Alu"  
repeat\_region  
16108..16144  
/rpt\_family="AT\_rich"  
repeat\_region  
16145..16309  
/rpt\_family="(TA)n"  
repeat\_region  
16310..16344  
/rpt\_family="AT\_rich"  
repeat\_region  
16814..16916  
/rpt\_family="MIR"  
repeat\_region  
16945..17421  
/rpt\_family="MaLR"  
repeat\_region  
17427..17488  
/rpt\_family="MIR"  
misc\_feature  
17507..17510  
/note="similar to Homo sapiens EST AA917407 (NID:g3057297)"  
misc\_feature  
17698..17701  
/note="similar to Homo sapiens EST AA224463 (NID:g1845081)"  
misc\_feature  
17878..18245  
/note="match to EST BF091143 (NID:g8481595)"  
repeat\_region  
18407..18491  
/rpt\_family="MIR"  
misc\_feature  
18456  
/note="similar to Homo sapiens EST A1078462 (NID:g3412670)"  
misc\_feature  
18456  
/note="similar to Homo sapiens EST BF966526 (NID:g11772035)"  
misc\_feature  
18536..19120  
/note="similar to Homo sapiens EST BE966526 (NID:g11772035)"  
misc\_feature  
18536..18918  
/note="similar to Homo sapiens EST A1078262 (NID:g3412670)"  
misc\_feature  
18537..19122  
/note="similar to Homo sapiens EST BG254743 (NID:g12764559)"  
misc\_feature  
18537..19080  
/note="similar to Homo sapiens EST AW007850 (NID:g5856628)"  
misc\_feature  
18537..18833  
/note="similar to Homo sapiens EST B1918363 (NID:g16182086)"  
misc\_feature  
18537..18548  
/note="similar to Homo sapiens EST AW157456 (NID:g6228857)"  
misc\_feature  
18543..19016  
/note="similar to Homo sapiens EST AW157456 (NID:g6228857)"

#### Alignment Scores:

| Pred. No.              | Score  | Length           |
|------------------------|--------|------------------|
| 1                      | 579.00 | 170227           |
| Percent Similarity:    | 62.45% | Matches: 136     |
| Best Local Similarity: | 55.51% | Conservative: 17 |
| Query Match:           | 25.91% | Mismatches: 32   |
|                        |        | Indels: 61       |

```

DB:
US-10-070-241b-1 (1-431) x AC093816 (1-170227)
QY 1 MetClnAlaLeuAsnIleThrProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20
Db 31301 ATGACGAGCGTTAAACATTACCGGGAGTATTTCTGAGTGTGAGAGAGAGAAATG 3142
QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40
Db 31241 ACAGCGGAGAGATGATCAATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 31182
QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60
Db 31181 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 31122
QY 61 PheClyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80
Db 31121 TTGGGAGATGCTTGTGTGTCTACGTTGGTACCGGAGAGAGAGAGAGAGAGAGAGAG 31062
QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100
Db 31061 AACATCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31002
QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGly 114
Db 31001 GTACCAATGCTTACAGAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 30942
QY 115 -----AlaPheIleCysLysMetValProPhe----- 123
Db 30941 CTTCGCTGCTTCTGCTTAAAGCTTGTGCTTTTAAAGAGATCTTCTCTCTCTCTCT 30882
QY 124 -----ValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140
Db 30881 CTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 30822
QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160
Db 30821 CTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 30784
QY 161 ArgAlaPheThrMetLeuGlyValValThrPheValAlaValIleValIleGlySerProMet 180
Db 30784 ----- 30784
QY 181 TrpHisValGlnGlnLeuGlnIleCysLysTyrAspPheLeuTyrGluLysGluHisIleCys 200
Db 30783 -----TCAATCGCTGCTGCTTCTTCTTA---GAACCTCCGAGCAGCTGCT 30745
QY 201 CysLeuGlnGlnTrpThrSerProValHisGlnIleTyrThrThrPheIleLeuVal 220
Db 30744 -----TCAATCGCTGAGATAACTTCTTTTGGCGGCTCTGCTGCTGCTGCTGCTGCT 30703
QY 221 IleLeuPheLeuLeu 225
Db 30702 CAGTGTCTGCTGCT 30688
RESULT 12
LOCUS AF268900
DEFINITION Rattus norvegicus neurone peptide FF receptor 2 (NPFF2) mRNA, complete cds.
ACCESSION AF268900
VERSION AF268900.1 GI:11907916
KEYWORDS Rattus norvegicus
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE Honini, J.A., Jones, K.A., Adham, N., Forray, C., Artymyshyn, R., Durkin, M.M., Smith, K.E., Tamm, J.A., Boteju, L.W., Yao, W.-J., Quan, Y., Ogozalek, K.L., Boyle, N., Kouranova, E.V., Vayssse, P.J., Wetzel, J.M., Branche, T.A., Gerald, C., and Borowsky, B. Identification and characterization of two G protein-coupled receptors for neuropeptide FF. J. Biol. Chem. 275 (50), 39324-39331 (2000)
AUTHORS
FEATURES
SOURCE 1..1254
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1254
/gene="NPFF2"
1..1254
/gene="NPFF2"
/feature="G protein-coupled receptor"
/codon_start=1
/product="neuropeptide FF receptor 2"
/protein_id="AAC41399.1"
/db_xref="GI:11907917"
/translational="MCKRWDSSGSDHMSGNDTQHPWYSDINITYMNYLYLQHPV
TAVTSSFLFFLCWGVNTVCFVIRNRYHIVTFINLAISDLLVGIEMPT
LDNITAGPFGSSCKISGLVSAVAVETLVAIVDFRCVYVFPKLTVKT
FYNIWGLAITIMPSAIMLHVQEKYRVRSLSHNKTSTVYVCHEDWPNQEMRI
YTVIATFVLAPELSIVIMYARICASLFTSAHSTQKLEQHWKSKKQVTKML
TVALPILSLPLWTLMMSDYADISPKLKVINYVPAHMLAFNCSSNP11YCF
ENENRSGFDQAEQCKVKVQEAAYGLRAKRLNDINTSGLLVHEPASQNSGELGC
RKSANDPTQESLMEETGENTST"
BASE COUNT 332 a 325 c 274 g 322 t
ORIGIN
Alignment Scores:
Pred No.: 1 89e-37 Length: 1254
Score: 508.00 Matches: 119
Percent Similarity: 51.07% Conservative: 72
Best Local Similarity: 31.82% Mismatches: 145
Query Match: 22.73% Indels: 48
DB: 10 Gaps: 9
ns-10-070-241b-1 (1-431) x AF268900 (1-1254)
QY 54 LeuIlePheAlaLeuAlaLeuPheClyAsnAlaLeuValPheTyrValValThrArgSer 73
Db 160 CAGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 219
QY 74 LysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeu 93
Db 220 AGGTACATCCACCGCTCCTTAATTTCTTCAATCTCAACCGGCAATTAAGACTTACTG 279
QY 94 IleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGly 113
Db 280 GTTGAATAATTTGTATGATGATATTAATTTGTATGATGATGATGATGATGATGATGAT 339
QY 114 GlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGluIle 133
Db 340 GGAACAGCATGCTGCAAGATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399
QY 134 LeuThrMetThrCysIleAlaValGluArgHisGlnGlyLeuValHisProPheLysMet 153
Db 400 TTCACCTTGGTGGCATTAGCGGATACACAGATTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
QY 154 LysTrpGlnTyrThrAsnArgAlaPheThrMetLeuGlyValValThrPheValAla 173
Db 460 AAG5-----CTACTGTGTAAGACACACTTTCTGATGATGATGATGATGATGATGATGAT 513

```

```

174 ValileValGlySerPro-----MetTrpHisValGlnGluLeuGluLeuLysTyr 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
514 ATCCACATATGATCCCACTGTCATCATCTGTCATGTCACAG-----GAGAAATATAC 567
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 AspPheLeuTyrGluLysGluHis-----IleCysCysLeuGluGlu 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
568 TACGGTGTGAGGTCAGCTCCCAATAAAACCCAGCACAGTCTACTGGTGTGGGAGGAT 627
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
205 TrpThrSerProValHisGlnLysIleTyrThrPheLeuValIleLeuPheLeu 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
628 TGGCCAAACAGAAATAGAGATGCTACACACCGGTGCTTTTCCACTATCTACCTG 687
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
225 LeuProLeuMetValMetLeuLeuTyrSerLysIleGlyTyrGluLeuTrpIleLys 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
688 GCTCCACTCTCCCTCATCTTATCATCTATGCAAGGATGGGGCTTCCCTCTCAAGACC 747
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 LysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGluMetSerLysIle 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
748 TCAGCACACAGCAGGTAAAG-----CAGAGGCTGGAGTASTGGCATATA 792
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 AlaArgLysLysLysArgAlaValIleMetMetValThrValAlaLeuPheAlaVal 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
793 TCCAGAGAAACAGAGGTATCAGATGCTGCTGACTGTGGCCCTCCTTTTCATCCTT 852
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
285 CysTrpAlaProPheHisValValHisMetIleGluTyrSerAsnPheGluLysGlu 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
853 TCCGTGGCTTCCCTGTGGACTCTGATGCTCTCAGACTATGCTCACCTGTCCACCTAAC 912
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
305 TyrAspValThrIleLysMetIlePheAlaIleValGlnIleIleGlyPheSerAsn 324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
913 AAAGTACGTGTGTCATCAATATTTATGCTACCTCTTTGCCACTGGCTCGCCCTTCGCAAT 972
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
325 SerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPhelLysLysValLeu 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
973 AGCAGTGTCAACCCCATCTATTTATGTTTATGTAATGAAATTTTGGCACTGTTTCCA 1032
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
345 SerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArgHisGlyAsn 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1033 GATGCTTTCAGTCTCTGCT-----CAAAAGAAAGTCAAAACCCAGCAAGCCTATGGC--- 1083
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
365 SerGlyIleThrMetLysLysAlaLysPheSerLeuArg----- 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1084 -----CTAGAGCTAAACGCACTGACATATACATATCTGGCTGTGTTG 1128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
380 -----GluAsnProValGluGluThrLysGly----GluAlaPheSer 392
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1129 GTCCATGAACCTGATCTCAAAACCCAGTGGGAAACCTTGGGATGTAGAAAAAGTSCA 1188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
393 AspGlyAsnIleGluValLysLeuCysGluGlnThrGluGlu 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1189 GACAATCCACACAGGAATCTTTGATGGAGGAACGGGACAA 1230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
AF268899 1263 bp mRNA linear PRI 20-DEC-2000
LOCUS
DEFINITION Homo sapiens neuropeptide FF receptor 2 (NPFF2) mRNA, complete cds.
ACCESSION AF268899
VERSION AF268899.1 GI:11907914
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1263)
Bonini, J.A., Jones, K.A., Adham, N., Forray, C., Artymyshyn, R.,
Durkin, M.M., Smith, K.E., Tamm, J.A., Boyle, L.W., Yao, W.-J.,
Quan, Y., Ogozalek, K.L., Boyle, N., Kouranova, E.V., Vaysse, P.J.,
Wetzel, J.M., Branchek, T.A., Gerald, C. and Borowsky, B.
Metzel, J.M., Branchek, T.A., Gerald, C. and Borowsky, B.
Identification and characterization of two G protein-coupled
receptors for neuropeptide FF
J. Biol. Chem. 275 (50), 39324-39331 (2000)
20564301

```

```

11024015
2 (bases 1 to 1263)
Bonini, J.A., Jones, K.A., Adham, N., Forray, C., Artymyshyn, R.,
Durkin, M.M., Smith, K.E., Tamm, J.A., Boyle, L.W., Yao, W.-J.,
Quan, Y., Ogozalek, K.L., Boyle, N., Kouranova, E.V., Vaysse, P.J.,
Wetzel, J.M., Branchek, T.A., Gerald, C. and Borowsky, B.
Direct Submission
Submitted (18-MAY-2000) Target Development and Assessment, Synaptic
Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652,
USA
FEATURES
source
location/Qualifiers
1..1263
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4q13.2-q13.3"
1..1263
/gene="NPFF2"
1..1263
/feature="NPFF2"
/feature="G protein-coupled receptor"
/codon_start=1
/product="neuropeptide FF receptor 2"
/protein_id="AAG41398.1"
/db_xref="GI:11907915"
/translation="MNEKWDNTSSNNWHPINWVNDTKHLYSDINITYVNYLHOPOV
AAFLISYFLIFFLCMGNTVCFIVMKNKMHVILNLIJLNLALSDLVGLHCPMPT
LLDNIAGHPFGNTCKISGLVQGISVAASVFTLVAIADRFQCVVYFFKPLTKIA
FVIMIIWIAITIMSPASMIHVOEEKYRVIRNSQNKISPVYVCRDNDNOMRKL
FTVITANITYLAPLSIVIMYGRIGISLPAAPHTGPKNLEWHVYSPKQKIRKML
LIVALLFLLSLYPLTLMMSYADLSPNELQIINITYIPFHLAFGNSVNPDIYVG
FFNEFRFGQFAFLQQLCKRAKMEAYALKAKSHVLINTSNQLVQESTFQNPFGT
LLYKSAEKPPQELVMEELKETNSSEI"
BASE COUNT 371 a 293 c 249 g 350 t
ORIGIN
Alignment Scores:
Pred. No. 6 78e-37 length 1263
Score 502.00 Matches 128
Percent Similarity 50.59% Conservative 87
Best Local Similarity 30.12% Mismatches 162
Query Match 22.46% Indels 48
EB 9 Gaps 12
US-10-070-241b-1 (1-431) x AF268899 (1-1263)
QY 14 LeuLeuArgAspHisAsnLeuThrArgGluGlnPheIleAlaLeuTyrTargLeuArgPro 33
Db 76 CTGACTCATATTAATTAATAC-----TATGTGAATCATATCTTCACAGCCT 126
QY 34 LeuValTyrThrProGluLeuProGlyArgAlaLysLeuAlaLeuValIleThrGlyVal 53
Db 127 CAAGTG-----GCAGCAATCTTCATATTTCCACATTT 159
QY 54 LeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValValThrArgSer 73
Db 160 CTGATCTCTTTTGTGCAATGCGGAAATACITGGTITGCTTTTATCTATATCAGGAAAC 219
QY 74 LysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeu 93
Db 220 AAAATATGTACACATCTNATATATTTTAAATCTTAAATCTTAAATCTTAAATCTTAA 279
QY 94 IleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGly 113
Db 280 GTTGGCATATCTGCGATGCGCTATACACATCTCTGGACAAATATATACAGGATGCGCATTT 339
QY 114 GlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGluLeu 133
Db 340 GGAACACAGATGTGCAAGATCAGTGGATGCGCCAGGAAATATCTGTCGACCTTCAGTC 399
QY 134 LeuThrMetThrCysIleAlaValGluArgHisGlnGlyLeuValHisProPheLysMet 153
Db 400 TTTACGTTACTTGTCAATTTGCTGTATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459

```





```

Db 460 AAG-----CTCACTATCAAGACAGCGTTTGTCATTATTATGATCATCTGGCTAGCC 513
QY 174 ValIleValGlySerPro-----MetTrpHisValGlnGlnLeuGluIleLysTyr 190
Db 514 ATCACTATATCTCTCCATCTGCACTATGTTACATGTCAGAA-----GAAATAATATAC 570
QY 191 AspPheLeuTyrGluLysGluHis-----TleCysCysLeuGluGluTrp 205
Db 571 CCACTGAGACCACTCCCAAGATAAACACAGTCCAGTCTACTGTCGCCGGAAGACTGG 630
QY 206 ThrSerProValHisGlnLysIleTyrThrThrPheIleLeuValIleLeuPheLeu 225
Db 631 CCAATCAGGAATAGAGAGATCTACACCACTGTGCTTTGGCAACATCTACCTGGCT 690
QY 226 ProLeuMetValMetIleLeuTyrSerLysIleGlyTyrGluLeuTrpIleLysLys 245
Db 691 CCCCTCTCCCTCATCTTCATCATGATGTCGAAGGATGGAATTCACACCTCTTC----- 741
QY 246 ArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGluMetSerLys----- 763
Db 742 -----AGGCGCTGAGTTCCTCACACA-----GCCAGGAAGAACAGGACCACTGC 786
QY 264 -----IleAlaArgGlyLysArgAlaValIleMetMetValThrValAlaLeu 281
Db 787 CAGTGTGTGTAGAGAGAGAGATATTAAAGAGCTGCTGATGTGGCCCTGCTT 846
QY 282 PheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsnPhe 301
Db 847 TTTATCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
QY 302 GluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGlnIleGly 321
Db 907 TCTCCAAATGAACCTCCAGATCATCAACATCTACATCTACCTCTTTGCACACTGCTGCCA 966
QY 322 PheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLysLys 341
Db 967 TTGGGCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
QY 342 AsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArg 361
Db 1027 GCTTTCCCAAGAGCTTTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCT 1086
QY 362 HisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGluAsn 381
Db 1087 TATGCC-----CTAAAGCTAAAGCCCATGTGCTCATAAAC 1122
QY 382 ProValGluGluThrLysGlyGluAla-----PheSerAspGlyAsnIleGluValLysLeu 400
Db 1123 ACATCTAATCAGCTTGTCCAGGAATCTACATTTCAAAACCTCATGCGGGAACCTTGCTT 1182
QY 401 CysGluGlnThrGluGluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGlu 420
Db 1183 TATAGGAAAGTGTCTGAAAGACCCCAACAGCAA-----TTAGTGCATGGAAGAA 1230
QY 421 LeuAlaGluAsnSer 425
Db 1231 TTAAGAAAGAACTACT 1245

RESULT 15
LOCUS AF257210
DEFINITION Homo sapiens G-protein coupled receptor HLWAR77 mRNA, complete cds.
ACCESSION AF257210
VERSION AF257210.1 GI:9309468
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

```

```

Tan, K.B., Shabon, D., Nuthallaganti, P., Wang, D.Y., Wilson, S.,
Bergsma, D.J. and Sarau, H.M.
Receptor for the pain modulatory neuropeptides FF and AF is an
orphan G-protein-coupled receptor
J. Biol. Chem. 275 (34), 25965-25971 (2000)
20408933
PUBMED 10851242
2 (bases 1 to 1560)
Elshourbagy, N.A., Ames, R.S., Fitzgerald, J.R., Foley, J.J.,
Chambers, J.K., Szekeres, P., Schmidt, D.B., Huckle, P.T., Dytko, G.M.,
Murdoch, P.R., Tan, K.B., Shabon, D., Nuthallaganti, P., Wang, D.Y.,
Wilson, S., Bergsma, D.J. and Sarau, H.M.
Direct Submission
Submitted (17-APR-2000) Molecular Biology, SmithKline Beecham
Pharmaceutical, 709 Swedeland Rd, P.O. Box 1539, King of Prussia,
PA 19406, USA
FEATURES
source
1..1560
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4q21"
48..1310
/notes="7-transmembrane receptor"
/codon_start=1
/product="G-protein coupled receptor HLWAR77"
/protein_id="AA87078.1"
/db_xref="GI:9409469"
/translation="MRRKNTNSSENHPPNNVNTKHHLSINITYVNYLHQPQV
AAFIISYELIFLCMGNTVCFIVFVNRKHHMTVNLFIINLAIISDLVIFCMPT
LVNIIAGMPFGNTMKISGLVHGISAASVFTLVIAVDFQCVVYFPRKLTIRKA
FYTIMIYVLAITIMSPAGMLHVEERYVYFRLNSQNTSPVYVWPEDFMKNEMRKI
YTVTLFANTYLAFLSLIVIMYGRIGISLFRAAVPHPTGKNOHVVSRKKQKLIKML
LIVALLLISLWPLMTLMLSDYADLSPELQLINLYIPFAHMLAFGNSVNPVIYV
FNENFRGFEQFQQLQCKRKPMPAYALAKSHVLINLINSQILVQHSFQNPQHGET
LLYRSKAEKFEQLWHEELKETTNSSEI"
BASE COUNT 476 a 287 c 287 g 450 t
ORIGIN
Alignment Scores:
Pred. No.: 8 64e-37 Length: 1560
Score: 502.00 Matches: 128
Percent Similarity: 50.59% Conservative: 87
Best Local Similarity: 30.12% Mismatches: 162
Query Match: 22.46% Indels: 48
DB: 9 Gaps: 12
US-10-070-241b-1 (1-43) x AF257210 (1-1560)
QY 14 LeuLeuArgAspHisAsnLeuThrArgGluGlnPheIleAlaLeuTyrArgLeuArgPro 33
Db 123 CTTACTCAGATATTAAATATACC-----TATGCAACTACTACTCTCCACAGCCT 173
QY 34 LeuValTyrThrProGluLeuProGlyArgAlaLysLeuAlaLeuValLeuThrGlyVal 53
Db 174 CAAGTG-----GAGAGAAATTTTATTTTACTTTT 206
QY 54 LeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValIleThrArgSer 73
Db 207 CTGATTTCTTTTCTGATGATGATAATGATGATGATGATGATGATGATGATGATGATGAT 266
QY 74 LysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeu 93
Db 267 AAACAIAIGACACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 326
QY 94 IleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGly 113
Db 327 GTTGGCAATCTGCAATGCTATATAAATATATATATATATATATATATATATATATATAT 386
QY 114 GlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValIleThrGluIle 133
Db 387 GGAAAAACGATGTGCAAAATATATGATGATGATGATGATGATGATGATGATGATGATGAT 446

```

```
QY 144 LeuThrMetThrCysIleAlaValGluArqIleGlnGlyLeuValHisProPheLysMet 153
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
447 TTACGTTAGTCTGAATTCCTGACATAGGTTCCAGTCTGGTCTACCCCTTTTAAACCA 506
QY 154 LysTrpGlnTyrThrAsnArgArgAlaPheThrMetLeuGlyValValTrrPheValAla 173
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
507 AAG-----CTCACTATCAAGNAGCGTTTGCTATTATTATGATCATCTGGGTCTAGCC 560
QY 174 ValIleValGlySerPro-----MetTrpHisValGlnGlnLeuGluIleLysTyr 190
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
561 ATCACTATATGCTCTCCATCTGCACTAATGTTACATGTCGAAGAA-----GAAATAATTAC 617
QY 191 AspPheLeuTyrGluLysGluHis-----IleCysCysLysLeuGluGluTrp 205
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
618 CCACTGACACCAACTCCACAAATAAAACCAGTCCAGCTACCTGCTGGGGGAAGACGG 677
QY 206 ThrSerProValHisGlnLysIleTyrThrThrPheIleLeuValIleLeuPheLeu 225
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
678 CCAATCAGGAATAATGACGAGATCTACACCACTGTGTGTTGGCAATCTACTAGTGGCT 737
QY 226 ProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrpIleLysLys 245
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
748 CCCCCTCCCTCAATTCATCATCATGATGGAAGATTGGAATTCACCTCTTC----- 788
QY 246 ArqValGlyAspGlySerValLeuArqThrIleHisGlyLysGluMetSerLys----- 263
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
789 -----AGGGCTGCAGTTCCTCACACA-----GCCAGGAAGAACCCAGGACGCTGC 833
QY 264 -----IleAlaArqLysLysArqAlaValIleMetMetValThrValValAlaLeu 281
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
844 CACGTGGTGTCAGGAAAGAGAGATATTAAATATGATGCTGTTGTTGTTGTTGTTGTT 893
QY 282 PheAlaValCysTrpAlaProPheHisValValHisMetMetIleGlyTyrSerAsnPhe 301
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
894 TTTATTCCTCATGCTGCCCTGCTGCACCTCTAATGATGCTCTCAGACTACGCTGACCTT 953
QY 302 GluLysGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnIleIleGly 321
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
954 TCTCCAAATGAACTGCAATACATAAATATCATACCTACCTCTTCTGACACTGCTGCGCA 1013
QY 322 PheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLysLys 341
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1014 TTCTGTAACGATGCTGTAATGATATATATATATATATATATATATATATATATATAT 1073
QY 342 AsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArg 361
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1074 GCTTTCACCAAGCTTTCAGCTCCAGCTCCGCAAAAGAACAGCAAGCCCTATCGAAGCT 1133
QY 362 HisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArqGluAsn 381
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1134 TATGCC-----CTAAAGCTAAAGCCATGCTGCTCATAAAC 1169
QY 382 ProValGluGluThrLysGlyGluAla---PheSerAspGlyAsnIleGluValLysLeu 400
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1170 ACATGTAATCAGCTTGTGCAAGATCTACATTTCAAAACCTCATGGGGAAACCTTGCTT 1229
QY 401 CysGluGlnThrGluGlnLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGlu 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1240 TATAGGAAAGCTGCTGAAACACCCCAACAGGAA-----TTAGTCAATGGAAGAA 1277
QY 421 LeuAlaGluAsnSer 425
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1278 TTAAGCAAACTACT 1292
```

Search completed: June 17, 2004, 07:59:47  
Job time : 4846 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2003, 05:04:42 : Search time 331 seconds  
(without alignments)  
2932 360 Million cell updates/sec

Title: us-10-070-241b-1

Perfect score: 2235

Sequence: 1 MCAINITPEQFSLRLPHNL : : RHLPRLSELAENSLPSLHH 431

Scoring table:

RUSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -P2N=xtb  
-O=/cgn2.1/uspro\_pool/US10070241/runat\_10062003\_095452\_20830/3FF-Query.fasta.1.583  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi  
-LIST=45 -LOCAL=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10070241 -ACGN\_1\_1\_263/runat\_10062003\_095452\_20830 -NCPG=6 -ICPG=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPRACK=100 -ICNGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database:

N\_Geneseq\_101002: \*  
1 /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT \*  
2 /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT \*  
3 /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT \*  
4 /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT \*  
5 /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT \*  
6 /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT \*  
7 /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT \*  
8 /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT \*  
9 /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT \*  
10 /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT \*  
11 /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT \*  
12 /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT \*  
13 /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT \*  
14 /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT \*  
15 /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT \*  
16 /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT \*  
17 /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT \*  
18 /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT \*  
19 /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT \*  
20 /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT \*  
21 /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT \*  
22 /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT \*  
23 /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT \*  
24 /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 2235  | 100.0       | 1293   | 22    | Human G protein-co |
| 2          | 2235  | 100.0       | 1296   | 22    | Human GTP-binding  |
| 3          | 2235  | 100.0       | 1296   | 22    | Human C protein-co |
| 4          | 2235  | 100.0       | 1296   | 22    | Human TGR346 cDNA  |
| 5          | 2235  | 100.0       | 1296   | 24    | Human neuroleptide |
| 6          | 2235  | 100.0       | 1710   | 24    | Human neuroleptide |
| 7          | 2235  | 100.0       | 2411   | 21    | cDNA encoding huma |
| 8          | 2235  | 100.0       | 2415   | 22    | Human G protein-co |
| 9          | 2232  | 99.9        | 1296   | 21    | Human orphan G pro |
| 10         | 2232  | 99.9        | 1296   | 21    | Human G protein co |
| 11         | 2226  | 99.6        | 1296   | 21    | Human G protein co |
| 12         | 1914  | 85.6        | 1737   | 24    | Mouse TGR346a cDNA |
| 13         | 1773  | 79.3        | 1772   | 24    | Mouse TGR346b cDNA |
| 14         | 1276  | 57.1        | 741    | 24    | Human DNA for pole |
| 15         | 863   | 38.6        | 511    | 21    | Human orphan G pro |
| 16         | 863   | 38.6        | 511    | 21    | Human G protein co |
| 17         | 711   | 31.8        | 432    | 22    | Human GPCR-like pr |
| c 18       | 711   | 31.8        | 432    | 24    | Human galatin rece |
| c 19       | 711   | 31.8        | 432    | 24    | Human DNA for pole |
| 20         | 508   | 22.7        | 1314   | 21    | Pat neuroleptide F |
| 21         | 502   | 22.5        | 1302   | 21    | Human neuroleptide |
| 22         | 502   | 22.5        | 1790   | 22    | Orphan receptor HG |
| 23         | 502   | 22.5        | 1824   | 22    | Human GPCR homolog |
| 24         | 502   | 22.5        | 1888   | 20    | Human 7-transmembr |
| 25         | 502   | 22.5        | 1888   | 21    | DNA encoding HLWAR |
| 26         | 502   | 22.5        | 1888   | 24    | Human cDNA encodin |
| 27         | 501.5 | 22.4        | 1440   | 21    | DNA encoding a hum |
| 28         | 501.5 | 22.4        | 1934   | 22    | AA509316           |
| 29         | 496   | 22.2        | 1263   | 21    | Orphan receptor HG |
| 30         | 486   | 21.7        | 1556   | 16    | Human NPY-Y7 recep |
| 31         | 477.5 | 21.4        | 1228   | 21    | Pat hippocampal ne |
| 32         | 474   | 21.2        | 1200   | 16    | Mouse NPY-Y7 recep |
| 33         | 469.5 | 21.0        | 1200   | 22    | Rat hippocampal ne |
| 34         | 469.5 | 21.0        | 1200   | 22    | Human neuroleptide |
| 35         | 469.5 | 21.0        | 1200   | 22    | Human neuroleptide |
| 36         | 469.5 | 21.0        | 1200   | 22    | Human neuroleptide |
| 37         | 469.5 | 21.0        | 1200   | 22    | Human neuroleptide |
| 38         | 469.5 | 21.0        | 1200   | 24    | Human neuroleptide |
| 39         | 468.5 | 21.0        | 3241   | 22    | Human neuroleptide |
| 40         | 468.5 | 21.0        | 2144   | 19    | Rhesus monkey neur |
| 41         | 464.5 | 20.8        | 1280   | 16    | AA095034           |
| 42         | 464.5 | 20.8        | 1335   | 23    | Human hippocampal  |
| 43         | 460.5 | 20.6        | 1335   | 23    | Nucleotide sequenc |
| 44         | 460.5 | 20.6        | 1633   | 20    | Non-endogenous hum |
| 45         | 446.5 | 20.0        | 1500   | 24    | Human 7 transmembr |
|            |       |             |        |       | NIY4-R gene. Mus   |

## ALIGNMENTS

RESULT 1  
AA095034  
ID AA095034 standard; cDNA; 1293 BP.  
XX AA095034  
XX AA095034  
XX AA095034  
DT 05 JUN 2001 (first entry)  
XX Human G protein-coupled receptor protein A027 encoding cDNA SEQ ID NO:2.  
DE Human G protein-coupled receptor, A027, non-specific, neuroleptide, e.  
KW hypotensive, antirheumatic, antiallergic, cardiac, antitussive,  
KW abortifacient, gene therapy; Alzheimer's disease, hypertension,  
KW pregnancy termination, rheumatism, allergy, angina pectoris; ss.  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers

```

FT CDS      1..1293
FT /*tag-a
FT /partial
FT /product= "AQ27"
FT /note= "G protein coupled receptor; no stop codon given"
XX
PN W0200116316-A1.
XX
PD 08 MAR 2001.
XX
XX 24-AUG-2000: 2000WO-JP05684.
XX
XX 27-AUG-1999: 99JP-0241530.
XX
XX (TAKE ) TAKEDA CHPM IND LTD.
XX
XX Watanabe T, Kikuchi K, Shintani Y;
XX
XX WPI: 2001 226689/23.
XX
XX P-PSDB: AAH74773.
XX
XX Human brain-originated guanosine triphosphate-protein-coupled receptor,
XX its salt and encoded gene, useful in (gene) diagnosis and development
XX of preventives and remedies for e.g. Alzheimer disease, hypertension
XX and rheumatism
XX
XX Claim 4: Page 78-79; 84pp: Japanese.
XX
XX The present sequence encodes a human G protein-coupled receptor
XX designated AQ27. AQ27 has mitogenic, neuroprotective, hypotensive,
XX antirheumatic, antiallergic, cardiant, antitumoral and abortifacient
XX activities. AQ27 can be used as a G protein-coupled receptor protein
XX agonist, G-protein-coupled receptor signal transducer and in gene
XX therapy. AQ27 can be used in the diagnosis and development of
XX preventives and remedies for diseases associated with dysfunction of
XX AQ27 e.g. Alzheimer's disease, hypertension, pre-eclampsia,
XX rheumatism, allergy and angina pectoris.
XX
XX Sequence 1293 BP; 342 A; 305 C; 301 G; 345 T; 0 other;
XX
XX Alignment Scores:
XX
XX Pred. No.:      2,86 244      Length:      1293
XX Score:          2235.00      Matches:      431
XX Percent Similarity: 100.00%      Conservative: 0
XX Best Local Similarity: 100.00%      Mismatches: 0
XX Query Match:      100.00%      Indels:      0
XX DB:               22          Gaps:        0
XX
XX US: 10 070 241B-1 (1-431) x AAH81818 (1-1293)
XX
XX QY 1 MetGlnAlaLeuAsnIleThrProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20
XX DB 1 ATGCGAGGGCTTAACATTAACCGGAGGAGATCTCTCGGCTCTCGGCGACACAAAGCTG 60
XX
XX QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40
XX DB 1 AGCGCGGAGCAGCTTCATCGCTCTGACCGGCTGGCAGCCGCTACACCCACAGCCTG 120
XX
XX QY 41 ProGlyArgAlaIleAlaLeuValIleuThrGlyValIleuIlePheAlaLeuAlaLeu 60
XX DB 1 CAGAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 180
XX
XX QY 61 PheGlyAsnAlaLeuValPheIleValIleThrArgSerIleAlaMetArgThrValThr 80
XX DB 1 TTTCGCAAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
XX
XX QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPheCysIlePro 100
XX DB 1 AACATCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
XX
XX QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaIlePheIleCysIleMet 120
XX DB 1 GTACACATGCTCCACAAACATTTCGACAAATTCGCTGGGGGTGCTTTTCATTTCGAAATG 360

```

```

QY 121 ValProPheValGlnSerThrAlaValValThrGluLeuThrMetThrCysIleAla 140
DB 361 GTGCCATTTCTCCAGCTTACCGCTGTGTGACAGAAATCTCACTATGACCTTCCATTCCT 420
QY 141 ValGluArgIleGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160
DB 421 GTGGAAAGGACCAAGGAAATTTGGTGTATTTTAAATGAAATGAGAAATAAATAAATAA 480
QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180
DB 481 AGGCTTTTCACTCAATGCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200
DB 541 TGGCAGCTGCAACAACCTTGAGATCAATATGACTTCTCTATATGAAAGGAAACACATCTGC 600
QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleTyrThrPheIleLeuVal 220
DB 601 TGTCTTAAGAGTGGAAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuLysIleGlyIleGly 240
DB 661 ATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260
DB 721 CTTTGCATAAAGAAACAGCTTGGGATGGTTCAGTCTCGAATATTCATCGAAAGAA 780
QY 261 MetSerLysIleAlaArgLysLysLysArgAlaValIleMetMetValThrValAla 280
DB 781 ATCTCCAAAATAGCCAGGAAGAAAGAGAGCTGTCTATTATGATGATGATGATGATGATGAT 840
QY 281 LeuPheAlaValCysTrpAlaProPheHisValIleHisMetMetIleGlyIleSerAsn 300
DB 841 CTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 400
QY 301 PheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGlnIle 320
DB 901 TTTCAAAAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 321 GlyPheSerAsnSerIleCysAsnProIleValIleValAlaPheMetAsnGluAsnPheLys 340
DB 961 GATTTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1020
QY 341 LysAsnValIleuSerAlaValCysTyrCysIleValIleAsnIleThrPheSerProAlaGln 360
DB 1021 AAAAAATGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380
DB 1081 AGCATGCAAAATTCAGGAATTCACAAATTCAGGAAGAAATTCAGGAAGAAATTCAGGAAG 1140
QY 381 AsnProValGluGluThrLysGlyGluAlaPheSerAspGlyAsnIleGluValLysLeu 400
DB 1141 AATCCAGTGGAGGAAACCAAGAGAAATTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 401 CysGlnGlnThrGluGluLysLysLysLysLysArgHisIleuAlaIleuPheArgSerGlu 420
DB 1201 TGTGAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431
DB 1261 CTGGCTGAGAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1293
XX
XX RESULT 2
XX AAH49502
XX ID AAH49502 standard; DNA: 1296 BP.
XX AC AAH49502;
XX DT 21-SEP-2001 (first entry)
XX

```

DE Human GTP-binding protein-coupled receptor GPRV11 coding sequence.  
 XX GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;  
 KW muscular; urinary; circulatory; anorectic; gene therapy; human;  
 KW guanosine triphosphate; G-protein; ds.  
 XX

OS Homo sapiens.

PN WO200148189-A1.

PD 05-JUL-2001.

XX 28-DEC-2000; 2000WO-JP09409.

XX 28-DEC-1999; 99JP-0375152.

PR 31-MAR-2000; 2000JP-0101339.

PR 23-MAY-2000; 2000JP-0155978.

XX (HELI-) HELIX RES INST.

PA Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;

PI Sugiyama T;

DR WPI; 2001-425663/45.

DR P-PSDB; AAG64289.

XX Family of guanosine triphosphate binding protein coupled receptors and  
 PT genes encoding them for treatment and prevention of diseases associated  
 PT with these receptors -

XX Claim 1; Pages 82-83; 137pp; Japanese.

XX The present sequence is the coding sequence for a human guanosine  
 CC triphosphate (GTP)-binding protein-coupled receptor. The receptor is  
 CC useful for the investigation, diagnosis, treatment and prevention of  
 CC diseases associated with GTP-binding protein-coupled receptors, including  
 CC neurological, circulatory, digestive system, immune system, muscle and  
 CC urinary system disorders. GTP-binding proteins are also known as  
 CC G-proteins.

XX Sequence 1296 BP; 344 A; 305 C; 301 G; 346 T; 0 other;

Alignment Scores:

Pred. No : 2,81e-243 Length: 1296  
 Score: 2235.00 Matches: 431  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-10-070-241b-1 (1-431) x AAH49502 (1-1296)

QY 1 MetGlnAlaLeuAsnIlePheProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20  
 DB 1 ATGCGAGCGCTTAACATTACCCGAGGAGAGTTCTCGGCTGCTCGGGGACCAACCTG 60  
 QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrProGluLeu 40  
 DB 61 ACGCGGGAGCAGTTATCGCTGTACCGGCTGCGGCGTGGACCGCTGTACACCCACAGCTG 120  
 QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
 DB 121 CCGGACGCGCCCAAGCTGCGCTCGTCTACCGCGTGTCTATCTTCGCGCTGCGCGCTC 180  
 QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80  
 DB 181 TTTGGCAATGCTCTGTGTCTTACGTGTGTGACCCGAGCAAGAGCATATCGCACCGTACCC 240  
 QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100  
 DB 241 AACATCTTTATCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysLysMet 120

DB 301 GTCCACATGCTCCAGAACATTTCCGACAACTGGCTGGGGGTGCTTTTCATTCAGAGATG 360  
 QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 DB 361 GTGCCATTGTCCAGTGTACGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
 QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnThrThrAsnArg 160  
 DB 421 GTGGAAGAGGACAGGAGCTTGTGCATCTTTTAAATGAAGTGGCAATACACCAACCGA 480  
 QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
 DB 481 AGGCTTTTCCAAATGCTAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 DB 541 TGGCAGCTGCAACAACCTTGAGATCAATATATCTGTATATGAAAGGAACACATCTGC 600  
 QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleTyrThrThrPheIleLeuVal 220  
 DB 601 TGCITAGAGAGTGGACCGCTGTGCCACCAAGATCTACACCACTTCATCCCTTGTTC 660  
 QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuLeuTyrSerLysIleGlyTyrGlu 240  
 DB 661 ATCCTCTCTCTCTGCTCTTATGGTGATGTTATTTCTGTACASTAAATGTTATGAA 720  
 QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260  
 DB 721 CTTTGGATAAAGAAAGAGTTGGGATGGTTCAGTGTCTGCAACTTTCATGGAAGAA 780  
 QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValValAla 280  
 DB 781 ATGTCCTCAATAGCCAGGAGAGAAAGAGCTGTCTATGATGCTGACAGTGTGGCT 840  
 QY 281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsn 300  
 DB 841 CTCTTTGCTGT 900  
 QY 301 PheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320  
 DB 901 TTTGAAAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 QY 321 GlyPheSerAsnSerIleCysAsnProIleValIleAlaPheMetAsnGluAsnPheLys 340  
 DB 961 GGATTTTCCAACTCCATCTCTAATCCCATTTGCTATGCAATTTATGAATCAAACTTCAA 1020  
 QY 341 LysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGln 360  
 DB 1021 AAAAAATGTTTGT 1080  
 QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380  
 DB 1081 AGGATGGAATTCAGGAATTTACATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
 QY 381 AsnProValGluGluThrLysGlyGluAlaPheSerAspGlyAsnIleGluValLysLeu 400  
 DB 1141 AATCCATGGAGGAGAAACCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
 QY 401 CysGluGlnThrGluGluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGlu 420  
 DB 1201 TGTGAACAGACAG 1260  
 QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
 DB 1261 CTGGCTGAGATTCCTCTTTAGACAGTGGGCAT 1293

RESULT 3

AAC87690

ID AAC87690 standard; DNA; 1296 BP.

XX

AC AAC87690;

XX

DT 27 MAR-2001 (first entry)  
 XX human G protein-coupled receptor AXOR16 DNA.  
 XX  
 KW AXOR16: human; G protein-coupled receptor; 7TM receptor;  
 KW chromosome 11q12.2; infection; viral; bacterial; fungal;  
 KW protozoan, HIV-1, HIV-2; pain, cancer, diabetes; obesity;  
 KW anorexia; bulimia; osteoporosis; asthma; allergy; urinary retention;  
 KW acute heart failure; hypotension; hypertension; angina pectoris;  
 KW myocardial infarction; stroke; ulcer; migraine; vomiting;  
 KW psychotic disorder, neurological disorder, anxiety, schizophrenia;  
 KW manic depression; bipolar disorder; depression; delirium; dementia;  
 KW severe mental retardation, dyskinesia; Parkinson's disease;  
 KW Huntington's disease; Gilles de la Tourette's syndrome; gene therapy;  
 KW vaccine; drug screening; signal transduction; ds.  
 XX  
 OS homo sapiens.  
 XX  
 PN W0200078809-A1.  
 XX  
 XX 28-DEC-2000.  
 XX  
 PF 19-JUN-2000; 2000WO-0516869.  
 XX  
 XX 21-JUN-1999; 99US-0337105.  
 XX  
 PA (SMK ) SMITHKLINE BEECHAM CORP.  
 XX  
 XX Elshourbaqy N, Shabon U;  
 XX W01; 2001-102699/11.  
 XX  
 DR P-PDH; AAH48963.  
 XX  
 XX New AXOR16, a G-protein coupled receptor polypeptide for screening  
 PT agonists and antagonists and for diagnosing and treating microbial  
 PT infections, cancer, neurological disorders and asthma  
 XX  
 PS Claim 2; Page 29; 31pp; English.  
 XX  
 CC The invention relates to the human G protein-coupled receptor AXOR16  
 CC (AAH48963) to DNA encoding AXOR16 (AAH48963), and to AXOR16 fragments  
 CC and variants. Like all G protein-coupled receptors, AXOR16 has 7  
 CC putative transmembrane domains and is involved in signal transduction.  
 CC AXOR16 has homology with G-protein coupled receptor polypeptide (NPYR1) F and  
 CC homology/ structural similarity with G-protein coupled receptor polypeptide Y (NPYR2A).  
 CC The human AXOR16 gene is located on chromosome 11q12.2. The invention  
 CC also relates to expression vectors and host cells comprising AXOR16 DNA,  
 CC to recombinant expression of AXOR16, and to an AXOR16-specific antibody  
 CC AXOR16 proteins and nucleotides may be used to treat a wide variety of  
 CC disorders including bacterial, fungal, protozoal and viral infections,  
 CC particularly HIV-1 or HIV-2 infections; pain; cancers; benign prostatic  
 CC hypertrophy; diabetes; obesity; anorexia; bulimia; osteoporosis; asthma;  
 CC allergies; urinary retention; acute heart failure; hypotension;  
 CC hypertension; angina pectoris; myocardial infarction; stroke; ulcers;  
 CC migraine, vomiting, psychotic and neurological disorders such as  
 CC anxiety, schizophrenia, manic depression, depression, delirium, dementia,  
 CC and severe mental retardation, and dyskinesias, such as Parkinson's  
 CC disease, Huntington's disease or Gilles de la Tourette's syndrome.  
 CC AXOR16 proteins, and nucleotides are useful as vaccines, and AXOR16  
 CC proteins, nucleotides and antibodies may be used in screening compounds  
 CC for their ability to modulate AXOR16 activity or expression. AXOR16  
 CC proteins are also useful for inducing an immunological response in a  
 CC mammal against the above diseases, for antibody production, and to  
 CC identify membrane bound or soluble receptors for AXOR16. AXOR16  
 CC nucleotides are also useful as diagnostic reagents and in chromosome  
 CC localisation and tissue expression studies. The present sequence  
 XX represents DNA encoding human AXOR16.

SQ Sequence 1296 BP; 344 A; 305 C; 301 G; 346 T; 0 other;

Alignment Scores:

Pred. No.: 2,810; 243 Length: 1296  
 Score: 2235; 00 Matches: 431

Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-10-070-241b-1 (1-431) x AAC87690 (1-1296)  
 QY 1 MetGlnAlaLeuAsnIleThrProGluGlnPheSerArgLeuValThrProGluLeu 20  
 Lb 1 ATGCAAGGCGCTTAACTTACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
 QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40  
 Lb 61 ACCGGGAGAGAGTTCATCGCTGCTACCGGCTCGAGCGCTCGCTCGCTCGCTCGCT 120  
 QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
 Lb 121 CCGGACCGCGCCAAAGCTGGCCCTCGGCTCACGAGAGAGAGAGAGAGAGAGAGAG 180  
 QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80  
 Lb 181 TTTGGTAATGCTCTGTGTTCTATAGTATGTAAGAGAGAGAGAGAGAGAGAGAGAG 240  
 QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100  
 Lb 241 AACATCTTTATCTGCTCCTTGGCGCTCAGTCACCTGCTCATCACCTTCTTGCATTCGC 300  
 QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMet 120  
 Lb 301 GTCACCATGCTCCAGAACATTTCCGACACCTGCTGGGGGCTGCTTTCATTTGCAAGATG 360  
 QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 Lb 361 GTGCCATTTGTCAGTCTACGGCTGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
 QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160  
 Lb 421 GTGAAAGGACCCAGGAGCTTGTGATCTTTTAAATGAAGTGGCAATACACCAACCCA 480  
 QY 161 ArgAlaPheThrMetLeuGlyValValThrLeuValAlaValIleValGlySerProMet 180  
 Lb 481 AGGCTTTTCAATAGCTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuGlyGlyGluHisIleCys 200  
 Lb 541 TGGCACTGTGAG 600  
 QY 201 CysLeuGluGluTrpThrSerProValHisSerLysIleTyrThrThrPheIleLeuVal 220  
 Lb 601 TGCATTACAGCAGTGGACAGCCCTGTGTCACAGACAGATCTACACCACTTCACTGCTGTC 660  
 QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGlu 240  
 Lb 661 ATCCCTCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260  
 Lb 721 CTTTGTATAAG 780  
 QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValAlaAla 280  
 Lb 781 ATGTCCAAAATAGCCAG 840  
 QY 281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsn 300  
 Lb 841 CTCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 301 PheGluLysGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320  
 Lb 901 TTTGAAAGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 QY 321 GlyThrSerAsnSerIleCysAsnProIleValTyrAlaIleMetAsnGluAsnPhenyls 340  
 Lb 961 TTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000

Db 961 GCATTTTCCAACTCCATCGTAATCCCAATGGCTATGCAATTAATCAATGAAACATTCAAA 1020  
 QY 341 LysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGln 360  
 Db 1021 AAAAAAGTGTGTCTGCAGTTTGTATTCATAGTAAATAAAGCTTCTCCAGCACAA 1080  
 QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380  
 Db 1081 AGGCATGAAATTCAGGAATTACAAATGATCGGGAAGCAAGCAAGTTTCCCTCAGAGAG 1140  
 QY 381 AsnProValGluGluThrLysGlyGluAlaPheSerAspGlyAsnIleGluValLysLeu 400  
 Db 1141 AATCCAGTGGAGGAACCAAGAGCAATTCAGTGTGATGCGCAACATTCGAAGTCAAAATG 1200  
 QY 401 CysGluGlnThrGluGluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGlu 420  
 Db 1201 TGTGACAGACAG 1260  
 QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
 Db 1261 CTGGCTGAGAAATTCCTCTTACAGAGTGGCAT 1293

## RESULT 4

ABK86287

ID ABK86287 standard; cDNA; 1296 bp.

XX AC ABK86287;

XX 27-AUG-2002 (first entry)

XX Human TGR346 cDNA.

XX Human; TGR346, gene, ss, G protein coupled receptor; GPCR; TGR;

XX TGR-associated disorder; signal transduction; renal failure, nephritis,

XX hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;

XX diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;

XX sleep disturbance; temperature regulation; blood pressure; hypohalimus;

XX circadian rhythm.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1-1296

XX /ftq= a

XX /product= "Human TGR346 protein"

XX W0200242458-A2.

XX 30-MAY-2002.

XX 21-NOV-2001; 2001WO-US43404.

XX 22-NOV-2000; 2000US-252841P.

XX 22-DEC-2000; 2000US-257636P.

XX 12-JAN-2001; 2001US-261377P.

XX 28-MAR-2001; 2001US-279554P.

XX 29-MAR-2001; 2001US-280696P.

XX (TUL- ) TULARIK INC.

XX Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;

XX WPI: 2002-463633/49.

XX P-PSDB; AA097735.

XX New isolated G-protein coupled receptor polypeptide, termed TGR, for

XX diagnosis and treatment of diseases such as renal failure, nephritis,

XX hypothyroidism, diabetes insipidus, and disturbances of thirst and

XX sleep -

XX Claim 42; Page 74; 98pp; English.

XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),

XX CC

CC termed TGR, and its associated nucleic acid. The sequences of the  
 CC invention are useful for identifying a compound that modulates signal  
 CC transduction and for identifying a mammal having a TGR-associated  
 CC disorder. The proteins and nucleic acids are useful in diagnosis and  
 CC treatment of diseases or conditions such as renal failure, nephritis,  
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,  
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,  
 CC sleep, temperature regulation, appetite, blood pressure or any other  
 CC syndrome or disease associated with the hypothalamus. The sequences can  
 CC be used in regulation of circadian rhythms, for use as genetic markers  
 CC for the identification of mutations associated with diseases resulting  
 CC from GPCR inactivation in particular cell types and for identification  
 CC of modulators of GPCR signal transduction. This sequence represents cDNA  
 CC encoding the human TGR346 polypeptide.

XX SQ Sequence 1296 BP; 344 A: 305 C: 301 G: 346 T: 0 other;

## Alignment Scores:

Score: 2416 2416 1296  
 Percent Similarity: 100.00% Matches: 431  
 Best Local Similarity: 100.00% Conservative: 0  
 Query Match: 100.00% Mismatches: 0  
 DB: 24 Indels: 0 Gaps: 0

US-10-070-241B-1 (1-431) x ABK86287 (1-1296)

QY 1 MetGlnAlaLeuAsnIleThrProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20  
 Db 1 ATCCAGCGCTTAACATACCCCGGAAATAGGCTCTCTGGGCTGCTGGGACACACACG 60  
 QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40  
 Db 61 ACGGCGACAGATTCATCGTCTGTATAGGCTCTCTGGGCTGCTGGGACACACACG 120  
 QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeu 60  
 Db 121 CCGGAGCGCCCAAGCTGGCCCTCGCTCGTCCACCGGCTGCTCATCTTCGCTTCC 180  
 QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80  
 Db 181 TTGGCAATGCTCTGTGTCTTCTAGCGGCGGCAATGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuIleThrPheCysIlePro 100  
 Db 241 AACATCTTATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysLysMet 120  
 Db 301 GTCACCATGCTCTGAGAAATATTCGCAACATGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 Db 361 GTCCCATTTGTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 QY 141 ValGluArgHisGlnIleLeuValHisProPheLysMetLysIleProIleThrAsnArg 160  
 Db 421 GTGAAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
 Db 481 AGGCTTTTCAATGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 Db 541 TGGCAGCTGCAACAACCTTGAGATCAAAATATGACTTCTTATATCAAAAGCAACACATCT 600  
 QY 201 CysLeuGluGluIleThrSerProValHisGlnIleLysIleCysIleThrThrPheIle 220  
 Db 601 TGTATAGAGAGATGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGlu 240





QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMet 120  
 DB 301 GTCCACATGCTCCCAACATTTCCGACAACTGGGCGGGGCGCTTCATTTCAATTCGCAATG 360  
 QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 DB 361 GTGGATTTTGTCAATTAAT 420  
 QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnThrThrAsnArg 160  
 DB 421 GTGGAAGGACACAGGAGCTGTGCTATCCCTTTTAAATGAAAGTGGCAATACACCAATGGA 480  
 QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
 DB 481 AGGGCTTTTCAACATGCTAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 540  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTrpAspPheLeuTrpGluLysGluHisIleCys 200  
 DB 541 TGTCAGGTGCAACATTTGAT 600  
 QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleValThrThrPheIleLeuVal 220  
 DB 601 TGCTTAGAGAGTGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 660  
 QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTrpSerLysIleGlyTrpGlu 240  
 DB 661 ATCCCTCTTCCCTCCCTCCCTCTTATGCTGATGCTTATTTCTGTACAGCTAAATTTGGTTATCAA 720  
 QY 241 LeuTrpIleLysLysArgValGlyAspCysValLeuArgThrIleHisGlyLysGlu 260  
 DB 721 CTTTGATTAAGAAAGAGTGGGGATGGTTCATGCTTCAATTTCAATTTCAATTTCAATTTCAAT 780  
 QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValIleAla 280  
 DB 781 ATGTCCAAATAGCCAGCAACAAACGACCTGTCATTTATGCTGTCACACTGCTGCT 840  
 QY 281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTrpSerAsn 300  
 DB 841 CTCTTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 301 PheGluLysGluTrpAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320  
 DB 901 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 QY 321 GlyPheSerAsnSerIleCysAsnProIleValTrpAlaPheMetAsnGluAsnPheLys 340  
 DB 961 GCATTTTCCAACTCCATCTGTAATCCCATTTCTGATGATGATGATGATGATGATGATGATGAT 1020  
 QY 341 LysAsnValLeuSerAlaValCysTrpCysIleValAsnLysThrPheSerProAlaGln 360  
 DB 1021 AAAAATGTTTGTCTGCAATTTGTTATTTGCAATTAATAAATAAATAAATAAATAAATAAATAA 1080  
 QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380  
 DB 1081 AGGCATGCAATTCAGGAATTAAT 1140  
 QY 381 AsnProValGluThrLysGlyGluAlaPheSerAspGlyAsnIleGluValLysLeu 400  
 DB 1141 AATCCAGTGGAGAAACCAAGAGGAGAGCAATTCAGTATGGCAACATTTGAAGTCAAAATTTG 1200  
 QY 401 CysGluGlnThrGluLysLysLysLysLysArgHisLeuAlaLeuPheArgSerGlu 420  
 DB 1201 TCTGAACAGACAG 1260  
 QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
 DB 1261 CTGGCTGAGATTTCTCTTTAGACAGTGGGCAI 1293

RESULT 6

ABK14957

ID ABK14957 standard; DNA; 1710 bp.

XX

AC ABK14957;

XX 08-MAY-2002 (first entry)  
 DT Human neuropeptide Y-like GPCR (NPY-like GPCR) DNA sequence.  
 DE  
 XX Human; neuropeptide Y-like G protein-coupled; receptor; antibody;  
 KW NPY-like GPCR; anorectic; antidiabetic; tranquilizer; hypertensive;  
 KW hypotensive; analgesic; cytostatic; antiasthmatic; osteopathic;  
 KW cardiovascular; antiulcer; nootropic; anticonvulsant; antiallergic;  
 KW neuroleptic; antidiabetic; antiparkinsonian; neuroprotective; gene therapy;  
 KW antidepressant; antiasthmatic; diuretic; chromosome 11q12.2; ds.  
 XX Homo sapiens.  
 US  
 XX Key Location/Qualifiers  
 PH misc\_feature 394..1689  
 FT /tag= a  
 FT /note= \*Human NPY-like GPCR gene sequence (ABK14956)\*  
 XX W02002-04-18 A2.  
 PN 17-JAN-2002.  
 XX 06-JUL-2001: 2601W0-US21276.  
 PF 06-JUL-2000: 2000US-216523P.  
 PR (PARR) HAYEK CORP.  
 PA Bloomquist BT, Zheleznik L;  
 PI WPI: 2002-148166/19.  
 DR  
 XX New neuropeptide Y-like G protein-coupled receptor, for detecting  
 PT regulators for preventing, ameliorating, or correcting e.g. obesity,  
 PT anorexia, diabetes, hypotension, hypertension, pain, cancers, or  
 XX neurological disorders -  
 PS Disclosure: Page 90; 91pp; English.  
 XX The present invention relates to a new polynucleotide encoding a  
 CC polypeptide neuropeptide Y-like G protein-coupled receptor (NPY-like  
 CC GPCR), comprising a 41 amino acid sequence, given in the specification.  
 CC The polynucleotide of the invention is used to produce a NPY-like GPCR  
 CC polypeptide. The complement of the polynucleotide can be used to detect  
 CC a coding sequence for a NPY-like GPCR. A reagent, preferably an antibody,  
 CC that binds to NPY-like GPCR can be used to detect it. The NPY-like GPCR  
 CC can be used to screen for agents that can regulate the activity of it. A  
 CC reagent that binds to or regulates NPY-like GPCR can be used to reduce  
 CC the biological activity of NPY-like GPCR, or treat a disorder associated  
 CC with abnormal expression of NPY-like GPCR. The reagent is useful for  
 CC preventing, ameliorating or correcting dysfunctions or diseases such as  
 CC obesity, anorexia, bulimia, diabetes, hypotension, hypertension, cocaine  
 CC withdrawal, pain, cancers, asthma, urinary retention, osteoporosis,  
 CC aquina pectoris, myocardial infarction, ulcers, allergies, congestive  
 CC heart failure, memory enhancement, cardiac and cerebral vasospasm,  
 CC pheochromocytoma, ganglioglioma, Huntington's disease, severe  
 CC mental retardation, Alzheimer's disease, Parkinson's disease, psychotic  
 CC and neurological disorders including anxiety, schizophrenia, manic  
 CC depression, delirium, dementia and dyskinesias. The polypeptide may also  
 CC be used in raising specific antibodies. The present nucleic acid sequence  
 CC contains the human NPY-like GPCR gene (ABK14956) located on chromosome  
 CC 11q12.2.  
 XX Sequence 1710 bp; 419 A; 439 C; 429 G; 423 T; 0 other;

## Alignment Scores:

| Pred No.               | 4 23e-243 | Length:       | 1710 |
|------------------------|-----------|---------------|------|
| Score:                 | 2235.00   | Matches:      | 431  |
| Percent Similarity:    | 100.00%   | Conservative: | 0    |
| Best Local Similarity: | 100.00%   | Mismatches:   | 0    |
| Query Match:           | 100.00%   | Indels:       | 0    |
| DB:                    | 24        | Gaps:         | 0    |



CC cytokine activity; cell proliferation; differentiation; immune  
 CC modulation; haematopoiesis regulation; tissue growth activity;  
 CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic  
 CC and thrombolytic activity; anti-inflammatory activity; and tumour  
 CC inhibition activity. The proteins may be administered to patients as  
 CC vaccines, and the nucleotides may be used as part of a gene therapy  
 CC regime. Diseases or conditions that may be treated using the proteins or  
 CC nucleotides of the invention include autoimmune diseases; genetic  
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,  
 CC fungal and viral infections, especially HIV; multiple sclerosis;  
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;  
 CC insulin dependent diabetes mellitus; and allergic reactions such as  
 CC asthma and anaemia. They may also be used for treating wounds, burns,  
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic  
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may  
 CC additionally be useful as contraceptives. Nucleic acid sequences of the  
 CC invention may be used in chromosome mapping, and as a source of  
 CC diagnostic primers and probes. The present sequence represents cDNA  
 CC encoding one of the 40 proteins of the invention.

XX  
 SQ Sequence 2411 BP; 656 A; 581 C; 569 G; 604 T; 1 other;

# Alignment Scores:

Pred. No.: 7,04e-243 Length: 2411  
 Score: 2235.00 Matches: 431  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-10-070-241B-1 (1-431) x AAA23435 (1-2411)

QY 1 MetGlnAlaLeuAsnIleThrProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20  
 DB ATGCAGGCGCTTACATTACCCCGACAGTCTCTCGGCTGCTCGCGGACCAACCTG 429  
 QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40  
 DB ACAGGAGGAGCAGTTCATCGCTCTACGCGCTGCGGACGCTGCTACCCAGAGCTG 489  
 QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuLeu 60  
 DB CCGGACCGGCAAGCTGCGCTCTACCGGCGCTGCTACCGGCGCTGCTCTCGGCGCTC 549  
 QY 61 PheGlyAsnAlaLeuValPheTyrValThrArgSerLysAlaMetArgThrValThr 80  
 DB TTTGGCAATGCTCTGGTCTTCTACGTGTGACCGCGCAGCGCATGCGCAGCTCACC 609  
 QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPheCysIlePro 100  
 DB AACATCTTTATCTGCTTGGGCTTACAGTGTGAGTGTGCTATCATCACTTCTTGTGATTC 669  
 QY 101 ValThrMetLeuGlnAsnIleSerAspAsnIlePheGlyGlyAlaPheIleCysLysMet 120  
 DB GTACCATGCTTCCAGAACATTTCCGACAACTTGGCTGGGGGTGCTTCAITTTGCAAGATG 729  
 QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 DB GTGCCATTGTTCAGTCTACCGCTGTTGTGACAAATCTCTCACTATGACCTGCAITGCT 789  
 QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160  
 DB GTGGAAGGGCACAGGAGCTTGTGATCAATCTTTTAAATGAAGTGGCAATACCAACCGA 849  
 QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
 DB AGGCGCTTTCACAAATGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 DB TGGCAGCTGCAACAACTTGTAGATCAATATGACTTCCTATATGAAAGGAACACATCTGC 969

QY 201 CysLeuGluGluThrPheSerProValHisGlnLysIleTyrThrThrPheIleLeuVal 220  
 DB TGTCTTGAAGAGTGGTACCTAGCCCTGTGTGTACCTAGAGATCTATACCACTTATCTCTTGT 1029  
 QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGlu 240  
 DB ATCTCTTCT 1089  
 QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260  
 DB CTCTTGATTAAGAAAAAGATTGGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149  
 QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValAla 280  
 DB ATGTCCTCAATAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1209  
 QY 281 LeuPheAlaValCysTrpAlaPropheHisValValHisMetMetIleGluTyrSerAsn 300  
 DB CTCTTGTCTGT 1269  
 QY 301 PheGluLysGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320  
 DB TTTGAAAAAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329  
 QY 321 GlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPhelys 340  
 DB GGATTTTCCAACTCCCATCTGTAATCCCATGCTCTATGCTATGCTATGCTATGCTATGCT 1389  
 QY 341 LysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGln 360  
 DB AAAAATCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449  
 QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysLysLysLysLysLys 380  
 DB AGGCATGGAAATTCAGGAATTCAGATGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1509  
 QY 381 AsnProValGluGluThrLysGlyGluAlaPheSerAspGlyAsnIleGluValLysLeu 400  
 DB AATCCAGTGGAGGAAACCAAG 1569  
 QY 401 CysGluGlnThrGluIleLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420  
 DB TGTGAACACACAG 1629  
 QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
 DB CTGGCTGAGAAATCTCTTTAGACAGTGGGCTAT 1662

RESULT 8  
 AAF81819  
 ID AAF81819 standard; cDNA; 2415 BP.  
 XX  
 AC AAF81819;  
 XX  
 DT 05-JUN-2001 (first entry)  
 XX Human G protein-coupled receptor protein A027 encoding cDNA.  
 DE  
 XX Human; G protein-coupled receptor; A027; nootropic; neuroprotective;  
 KW hypotensive; antirheumatic; antiallergic; cardiant; antianigmal;  
 KW abortifacient; gene therapy; Alzheimer's disease; hypertension;  
 KW pregnancy termination; rheumatism; allergy; angina pectoris; ss.  
 XX Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 354..1649  
 FT /\*tag= a  
 FT /product= "A027"  
 FT /note= "G protein coupled receptor"  
 XX  
 PN W0200116316-A1.



```

FT      CDS
FT      1..1296
FT      /*tag= a
FT      /product= "hRUP4"
FT      /note= "human orphan G protein-coupled receptor"
XX      WO200031258-A2.
XX      PD
XX      02-JUN-2000.
XX      PF      13-OCT-1999; 99WO-US23687.
XX      PR      20-NOV-1998; 98US-0109213.
XX      PR      16-FEB-1999; 99US-0120416.
XX      PR      26-FEB-1999; 99US-0121852.
XX      PR      12-MAR-1999; 99US-0123946.
XX      PR      12-MAR-1999; 99US-0123949.
XX      PR      28-MAY-1999; 99US-0136436.
XX      PR      28-MAY-1999; 99US-0136437.
XX      PR      28-MAY-1999; 99US-0136439.
XX      PR      28-MAY-1999; 99US-0136567.
XX      PR      28-MAY-1999; 99US-0137127.
XX      PR      28-MAY-1999; 99US-0137131.
XX      PR      29-JUN-1999; 99US-0141448.
XX      PR      29-SEP-1999; 99US-0156555.
XX      PR      29-SEP-1999; 99US-0156633.
XX      PR      29-SEP-1999; 99US-0156634.
XX      PR      29-SEP-1999; 99US-0156653.
XX      PR      01-OCT-1999; 99US-0157280.
XX      PR      01-OCT-1999; 99US-0157281.
XX      PR      01-OCT-1999; 99US-0157282.
XX      PR      01-OCT-1999; 99US-0157293.
XX      PR      01-OCT-1999; 99US-0157294.
XX      PR      12-OCT-1999; 99US-0416760.
XX      PR      12-OCT-1999; 99US-0417044.
XX      PA      (AREN-) ARENA PHARM INC.
XX      PI      Chen R, Dang HT, Liaw CW, Lin I;
XX      WPI: 2000-400068/34.
XX      DR      P-PSDB; AAV71309.
XX      PT      Novel human orphan G protein-coupled receptors and the encoding cDNAs
XX      PT      for use in the identification of G protein coupled receptor agonists -
XX      PS      Claim 73, Page 88-89, 102pp; English.
XX      CC      The present sequence is a cDNA encoding hRUP4, an endogenous human
XX      CC      orphan G protein-coupled receptor (GPCR). The full length hRUP4 cDNA was
XX      CC      cloned by RT-PCR with human brain cDNA as template. The hRUP4 PCR
XX      CC      fragment obtained was an alternatively spliced form of the EST (expressed
XX      CC      sequence tag) clone AF07658. The orphan GPCR of the invention, like
XX      CC      all GPCRs has seven trans-membrane alpha helices with an extracellular
XX      CC      N-terminus and an intracellular C-terminus. However, no endogenous
XX      CC      ligands has yet been identified for the proteins of the invention. The
XX      CC      orphan GPCRs may be used in the identification of their endogenous
XX      CC      ligands, and to screen potential GPCR agonists and antagonists for use as
XX      CC      pharmaceutical agents. The proteins may also be used in the study of
XX      CC      GPCR-mediated signalling cascades, and to elucidate their precise role in
XX      CC      normal and diseased human conditions. Nucleic acid encoding human orphan
XX      CC      GPCRs may be used for tissue localisation expression analysis to provide
XX      CC      information about their function in healthy and pathological states.
XX      SQ      Sequence 1296 BP; 344 A; 304 C; 302 G; 346 T; 0 other;
XX      Alignment Scores:
XX      Prd No.: 6, 150-243 Length: 1296
XX      Score: 2212.00 Matches: 430
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 99.77% Mismatches: 0
XX      Query Match: 99.87% Indels: 0
XX      DB: 21 Gaps: 0

```

```

US-10-070-241b-1 (1-431) x AAD01136 (1-1296)
QY      1 MetGlnAlaLeuAsnIleThrProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20
DB      1 ATGACAGGCGGTAAATTATACACGAGAGAGATCTCTTGGGTGGTGGAGAAATATG 60
QY      21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40
DB      61 ACGGGAGGAGGAGTTTCATCGCTTATACGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
QY      41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60
DB      121 CCGGGACGCGGCAAGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY      61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80
DB      181 TTTGGTAATGCTTGGTGTATAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
QY      81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100
DB      241 AATATTTATCTGCTCTTGGGAGTACAGTGAAGTGTATATAGTGTGTGTGTGTGT 400
QY      101 ValThrMetLeuGlnAsnIleSerAspAsnIlePheGlyAlaPheIleCysLysMet 120
DB      301 GTCACTCATGCTTCAAGAAATATTCGCAAACTGGGCGGCGGCGGCGGCGGCGGCGG 360
QY      121 ValProPheValGlnSerThrAlaValValThrIleIleLeuThrMetThrCysIleAla 140
DB      361 GTGCAATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY      141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysIleIleThrAsnArg 160
DB      421 GTGCAAAAGCACTCAAGGAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY      161 ArgAlaPheThrMetLeuGlyValValThrLeuValAlaValIleValGlySerProMet 180
DB      481 ACGGCTTTCACAAATGCTAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY      181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200
DB      541 TGGCACTGTGCAAACTTGGATCAATATGATGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY      201 CysLeuGluGluIleGlyThrSerProValHisLysLysIleIleIleThrPheLeuVal 220
DB      601 TGTATTAAAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY      221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleIleIle 240
DB      661 AAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY      241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260
DB      721 CTTTGTATAAAGAAAGAGTTGGGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY      261 MetSerLysIleAlaArgLysLysLysArgAlaValIleMetMetValThrValAla 280
DB      781 ATGTAAATAATAGTAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY      281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsn 300
DB      841 CTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY      301 PheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320
DB      901 TTTGAAAAAGGAATATGATGATGTCAATTAAGATGATTTTGTGTGTGTGTGTGTGTGT 960
QY      321 GlyPheSerAsnSerIleCysAsnProIleValIleTrpAlaPheMetAsnGluAsnIleCys 340
DB      961 GATTTTTCATCTCATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
QY      341 LysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGln 360
DB      1021 AAAAAATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080

```

QY 361 ArqHisGlyAsnSerGlyIleThrMetMetArqGlyLysAlaLysPheSerLeuArgGlu 380  
 DB 1081 AGGATGGAATTTAGAGATTACATGATGAGGAGAAAGAAAGTTTACTTACAGAG 1140  
 QY 381 AsnProValGluGluThrLysGlyGluAlaPheSerAspGlyAsnIleGluValLysLeu 400  
 DB 1141 AATCAGTGGAGAAACCAAGGAGACAGCATTCAGTCATGCGCAACATTTGAAGTCAAAATTG 1200  
 QY 401 CysGluGlnThrGluGluLysLysLysLeuLysLeuLysLeuLysLeuLysLeu 420  
 DB 1201 TCTCAACACAG 1460  
 QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
 DB 1261 CTGGCTGAGAAATTCCTTTAGACAGTGGGCAT 1293  
 RESULT 10  
 ID AAA46047  
 XX AAA46047 standard; cDNA: 1296 bp.  
 AC AAA46047;  
 DT 22-AUG-2000 (first entry)  
 XX Human G protein coupled receptor hRUP4 encoding cDNA SEQ ID NO:39.  
 DE Human: G protein coupled receptor; GPCR; transmembrane receptor;  
 KW identification; agonist; screening; therapeutic; pharmaceutical;  
 KW mutant; ss.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 XX WO200022141-A2.  
 PN 20-APR-2000.  
 PP 13-OCT-1999; 99WO-US24065.  
 XX 13 OCT 1998; 980S-0170496.  
 PR 12 NOV 1998; 980S-0108027.  
 PR 20 NOV 1998; 980S-0109214.  
 PR 27 NOV 1998; 980S-0110060.  
 PR 16 FEB 1999; 990S-0120416.  
 PR 26 FEB 1999; 990S-0121852.  
 PR 12 MAR 1999; 990S-0124944.  
 PR 12 MAR 1999; 990S-0124945.  
 PR 12 MAR 1999; 990S-0124946.  
 PR 12 MAR 1999; 990S-0124948.  
 PR 12 MAR 1999; 990S-0124949.  
 PR 12 MAR 1999; 990S-0124951.  
 PR 28 MAY 1999; 990S-0136446.  
 PR 28 MAY 1999; 990S-0136447.  
 PR 28 MAY 1999; 990S-0136449.  
 PR 28 MAY 1999; 990S-0137127.  
 PR 28 MAY 1999; 990S-0137131.  
 PR 30 JUN 1999; 990S-0141448.  
 PR 27 AUG 1999; 990S-0151114.  
 PR 04 SEP 1999; 990S-0152524.  
 PR 29 SEP 1999; 990S-0156634.  
 PR 29 SEP 1999; 990S-0156555.  
 PR 29 SEP 1999; 990S-0156644.  
 XX (AREN-) ARENA PHARM INC.  
 XX Rehan DP, Lehmann-Bruinsma K, Chaimers DL, Chen K, Dang HT;  
 PI Gore M, Liaw CW, Lin L, Lowitz K, White C;  
 XX WPI: 2000-117986/27.  
 DR P-PSDB: AAB02843.  
 XX Non-endogenous, human G protein-coupled receptors for screening

PT receptor, inverse or partial agonists useful as therapeutic agents  
 XX Example 1: Page 118-119; 187pp; English.  
 CC The present invention describes transmembrane receptors, preferably  
 CC human G protein coupled receptors (GPCR), for which the endogenous  
 CC ligand is unknown (orphan GPCR receptors). More specifically the present  
 CC invention relates to non-endogenous, constitutively activated versions  
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
 CC the direct identification of candidate compounds as receptors agonists,  
 CC inverse agonists or partial agonists for use as pharmaceutical agents.  
 CC AAA46047 to AAA46126 and AAA46825 to AAA46859 represent sequences used in  
 CC the exemplification of the present invention.  
 XX Sequence 1296 BP; 344 A; 304 C; 302 G; 346 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 6,15e-243 Length: 1296  
 Score: 2232.00 Matches: 430  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.77% Mismatches: 0  
 Query Match: 99.87% Indels: 0  
 DB: 21 Gaps: 0  
 US-10-070-241B-1 (1-431) x AAA46037 (1-1296)  
 QY 1 MetGlnAlaLeuAsnIleThrProGluGlnPheSerArqLeuLeuArqAspHisAsnLeu 20  
 DB 1 ATGCAAGAGGTTAAATTAAATTAATGAGGAGAAATTTCTGGATGATGAGGAGAGAG 60  
 QY 21 ThrArgGluGlnPheIleAlaLeuTyrArqLeuArqProLeuValTyrThrProGluLeu 40  
 DB 61 ACAGGAGAGAGAGTTCATGCTTCTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 QY 41 ProGlyArqAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
 DB 121 CCGGAG 180  
 QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArqSerLysAlaMetArqThrValThr 80  
 DB 181 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
 QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuIleThrPhePheCysIlePro 100  
 DB 241 AACATCTTTATCT 300  
 QY 101 ValThrMetLeuGlnAsnIleSerAspSerTrpLeuGlyGlyAlaPheIleCysLysMet 120  
 DB 301 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360  
 QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 DB 361 GTGTCATTTGTCCAGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160  
 DB 421 GTGTAAG 480  
 QY 161 ArgAlaPheThrMetLeuGlyValValThrLeuValAlaValIleValGlySerProMet 180  
 DB 481 AGGGCTTTCACAAATGCTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 DB 541 TGAATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600  
 QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleTyrThrThrPheIleLeuVal 220  
 DB 601 TGTATTAAAG 660  
 QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGlu 240  
 DB 661 AAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720







1 ATGAGGCGCTCAACATCACCGCGGAGCAGTTTTTCCCGGCTGCTGAGCGGCACCAACCTG 60  
 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArqProLeuValTyrThrProGluLeu 40  
 61 ACTGGGAACAGTTCATTCATCGTATGAGCTGAGAGAGTGTGCTGATACATGCTGAGTGTG 120  
 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
 121 CCGGAGAGAGTAAATGAGGCTGAGTGAAGATCATTTTTGGCTGAGGCTG 180  
 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80  
 181 TTTGGCAACTCTCTGGTCAATCATGCTGACCGCCAGCAAGGCGCATCGCGCACCGTCAAC 240  
 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100  
 241 AAATTCATCTGTTCTGAGATTCATGATGCTGATGCTGATGCTGATGCTGATGCTG 300  
 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysIleMet 120  
 301 GTCAGGATGCTCCAGAACATCTCCAGCAAGTGGCTGGTGGTGGTGGTGGTGGTGGTGG 360  
 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 361 GTGGCCTTCGTCAGTCCACTGCTGTTGTGACGGAAATCCCTCACCATCATGTCATGCT 420  
 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArg 160  
 421 GTTGAGAGGCACCAAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
 481 AGGGCTTTTCAAACTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540  
 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 541 TGGCAGGTACAAACCCCTCGAGATTAAAGTATGATCTCTCTATGAGAAAGAACATGCTG 600  
 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleTyrThrPheIleLeuVal 220  
 601 TGTGTGAAGAGTGGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAG 660  
 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGlu 240  
 661 ATCTCTCTCCCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 720  
 241 LeuTrpIleLysLysArgValLysAspLysSerValLeuArgThrIleHisLysLysIle 260  
 721 CTGTGGATCAAGAGAGAGATGGAGACAGTTTACACATTCACATTCACACCGGAAAGAA 780  
 261 MetSerLysIleAlaArgLysLysLysArgAlaValIleMetMetValThrValValAla 280  
 781 ATGTCCAAATAGCCAGGAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
 281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsn 300  
 841 CTCTTCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 900  
 301 PheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleValGluIleIle 320  
 901 TTTGAAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
 321 GlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLys 340  
 961 GGTCTTTTCACTCCATCTGTAATGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1020  
 341 LysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGln 360  
 1021 AAGAAATTTTGTCTGGGTTTGTATTGTCATAGTAGAGAAACCTTCTCCCGAGGACAG 1080  
 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380  
 1081 AAGCTGCAAAATCTTGGCATTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1140

QY 381 AsnProValGluCluThrLysGlyGluAlaPheSerAspClyAsnIleGluValLysIleu 400  
 Db 1141 CGTCCATGGCGGAGAGCAAAAGAGAGACTTATTACAGGATGGCAAGCTTGTATGCTCAAAATG 1200  
 QY 401 CysGluGlnThrGluCluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGlu 420  
 Db 1201 TGTGAGTGAAGTCAAGGAGCAAAAGAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1260  
 QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431  
 Db 1261 CTCTTCGCAAAACCTCTACTTTGGCGACGTGGACAT 1293  
 RESULT 13  
 ABK86293  
 ID ABK86293 standard; cDNA; 1772 BP.  
 XX  
 AC ABK86293;  
 DT 27-AUG-2002 (first entry)  
 DE Mouse TGR346b cDNA.  
 XX  
 KW Mouse; TGR346b; gene; ss; G-protein coupled receptor; GPCR; TGR;  
 KW TGR-associated disorder; signal transduction; renal failure; nephritis;  
 KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;  
 KW diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;  
 KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;  
 KW circadian rhythm.  
 XX  
 OS Mus sp.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..1251  
 FT /\*tag= a  
 FT /product= "Mouse TGR346b protein"  
 XX  
 PN W0200242458-A2.  
 XX  
 XX 30-MAY-2002.  
 XX  
 PF 21-NOV-2001; 2001WO-US43404.  
 XX  
 PE 22-NOV-2000; 2000US-252841P.  
 PE 22-DEC-2000; 2000US-257636P.  
 PE 12-JAN-2001; 2001US-261377P.  
 PE 28-MAR-2001; 2001US-279554P.  
 PE 29-MAR-2001; 2001US-280696P.  
 XX  
 PA (TULA-) TULARIK INC.  
 XX  
 XX Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;  
 PI P-PSDB; AAU97741.  
 DR WPI; 2002-463633/49.  
 DR  
 XX  
 XX New isolated G-protein coupled receptor polypeptide, termed TGR, for  
 PT diagnosis and treatment of diseases such as renal failure, nephritis,  
 PT hypothyroidism, diabetes insipidus, and disturbances of thirst and  
 PT sleep.  
 XX  
 PS Claim 22; Page 78-79; 98pp; English.  
 XX  
 XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),  
 CC termed TGR, and its associated nucleic acid. The sequences of the  
 CC invention are useful for identifying a compound that modulates signal  
 CC transduction and for identifying a mammal having a TGR-associated  
 CC disorder. The proteins and nucleic acids are useful in diagnosis and  
 CC treatment of diseases or conditions such as renal failure, nephritis,  
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,  
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,  
 CC sleep, temperature regulation, appetite, blood pressure or any other  
 CC syndrome or disease associated with the hypothalamus. The sequences can

CC be used in regulation of circadian rhythms, for use as genetic markers  
 CC for the identification of mutations associated with diseases resulting  
 CC from GPCR inactivation in particular cell types and for identification of  
 CC modulators of GPCR signal transduction. This sequence represents cDNA  
 CC encoding the mouse Tr345b polypeptide.

XX Sequence 1772 BP; 494 A; 390 C; 403 G; 485 T; 0 other;

Alignment Scores:  
 Pred. No.: 111e-190 Length: 1772  
 Score: 1773.00 Matches: 343  
 Percent Similarity: 89.44% Conservative: 38  
 Best Local Similarity: 80.52% Mismatches: 42  
 Query Match: 79.33% Indels: 4  
 DB: 24 Gaps: 2

US 10 070-241B-1 (1-41) x ABR86294 (1-1772)

QY 5 AsnIleThrProGluGluPheSerArgLeuLeuLeuArgAspHisAsnIleThrArgGluGln 24  
 DB 10 AACTTGACGGGGAGCAGCTCTCGGGCTGTTCGGCTGACAACTGACGAGGCTCAG 69  
 QY 25 PheIleAlaLeuTyrAcuLeuArgProLeuValTyrThrProGluLeuProGlyArgAla 44  
 DB 70 TTCAATGGGGAATATGGGCTGGGGTCACTGGTGTCTACCCGGAGCTTCCCGCGCGCC 129  
 QY 45 LysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeuPheGlyAsnAla 64  
 DB 130 AAGGTGGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 189  
 QY 65 LeuValPheTyrValValThrArgSerLysAlaMetArgThrValThrAsnIlePheIle 84  
 DB 140 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 249  
 QY 85 CysSerLeuAlaLeuSerAspLeuIleThrPhePheCysIleProValThrMetLeu 104  
 DB 250 TCGTCCCTGGGACATCAGGACGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309  
 QY 105 GluAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysLysMetValProPheVal 124  
 DB 310 CAGAGGCTGCTGGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369  
 QY 125 GluSerThrAlaValValThrGluIleLeuThrMetThrCysIleAlaValGluArgHis 144  
 DB 370 CAGTATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 429  
 QY 145 GluGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArgAlaPheThr 164  
 DB 440 CAGGACCTTCCTCCATCCCTTTAAAAAGGCGGAGTACACAAATCAAGAGCTTTTCACA 489  
 QY 165 MetLeuGlyValValTrpLeuValAlaValIleValGlySerProMetTrpHisValGln 184  
 DB 490 ATGCTAGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 549  
 QY 185 GluLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCysCysLeuGluGlu 204  
 DB 550 GCACTTGAGATTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609  
 QY 205 TrpThrSerProValHisGlnLysIleTyrThrPheIleLeuValIleLeuPheLeu 224  
 DB 610 TGGAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 669  
 QY 225 LeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrpIleLys 244  
 DB 670 CTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729  
 QY 245 LysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGluMetSerLysIle 264  
 DB 740 AAAAGAAATCGGGGATGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789  
 QY 265 AlaArgLysLysLysArgAlaValIleMetMetValThrValValAlaLeuPheAlaVal 284  
 DB 790 GTCACAAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849

QY 285 CysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsnPheGluLysGlu 304  
 DB 850 IGITGGGCACTTTTCAATCGGTTCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 909  
 QY 305 TyrAspAspValThrIleLysMetIlePheAlaIleValIleIleGlyPheSerAsn 324  
 DB 910 TATGATCAACTCAATCAAGATGATTTTTCGTATAGTCAAAATATATGATTTTTCAC 969  
 QY 325 SerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLysLysValLeu 344  
 DB 970 TCCATCTGTAATCCCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029  
 QY 345 SerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArgHisGlyAsn 364  
 DB 1030 TCTCGGTTTGTATTTCATTGTAAGAAATATTTTTCATGTAAGAAATGTAAGAAAT 1089  
 QY 365 SerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGluAsnProValGlu 384  
 DB 1090 TCAGGAGCTATGCTGATGCTGACAGGCGCAAGATAGCTGCAAGACAGACAAATGCTGTA 1146  
 QY 385 GluThrLysGlyGluAlaPheSerAspLysAsnIleLeuValLysLeuCysLysIleThr 404  
 DB 1147 GATATCAAAAGCAAACTATTTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206  
 QY 405 GluGluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGluLeuAlaGln 423  
 DB 1207 GAAAG 1265  
 QY 423 uAsnSerProLeuAsp 428  
 DB 1266 GACCTCTGCTGATGAC 1281  
 RESULT 14  
 AAS98050  
 ID AAS98050 standard; DNA; 741 BP.  
 XX AAS98050;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Human DNA for potential G protein-coupled receptor #8.  
 KW Human; G protein-coupled receptor; GPCR; ds; GAI4; galanin receptor;  
 KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;  
 KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;  
 KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;  
 KW depression; epilepsy; macular degeneration; lymphoma; melanoma;  
 KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;  
 KW psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;  
 KW tuberculosis; cognition disorder; memory disorder; anorexia;  
 KW hormonal release disorder; cardiovascular activity disorder;  
 KW pain perception disorder; obesity; diabetes; diabetes; obesity;  
 KW diabetes; hyperlipidaemia; stroke; gene therapy.  
 KW Homo sapiens.  
 QS  
 XX W0200185791-A1  
 PN  
 XX 15-NOV-2001.  
 PD  
 XX  
 PF 11-MAY-2001; 2001WO-US15332.  
 XX  
 PP 11-MAY-2000; 2000US-204217P  
 PP 18-MAY-2000; 2000US-205945P.  
 XX  
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.  
 PA  
 XX Brown JP, Miller M, Burner G, Fabre-Suver C, Pritchard D;  
 PI WPI; 2002-066595/09.  
 XX  
 XX Novel G protein-coupled receptor polypeptides including galanin

receptor polypeptides useful for identifying modulators that are useful for treating Alzheimer's disease, psoriasis, melanoma, multiple sclerosis, stroke

Disclosure: Page 117; 144pp; English.

The invention relates to an isolated polypeptide encoded by a nucleic acid molecule that is at least 80% identical to the G protein-coupled (GPCR) polynucleotides included in the specification. Also included are probes based on the GPCR sequences (including antisense probes), a host cell comprising an expression vector comprising the GPCR sequence, antibodies raised against the polypeptides, and methods of identifying modulators of the polypeptides. The polypeptides are useful for identifying modulator compounds which function as modulators, activators, repressors, agonists or antagonists of the novel GPCR polypeptides including the GAL4 polypeptide. The antibodies and nucleic acid probes as described above can be used to detect the presence of the polypeptides and nucleic acids and are used to diagnose a variety of diseases or disorders in which GPCRs are involved e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma, atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy, chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease, depression, epilepsy, macular degeneration, lymphoma, melanoma, multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease, psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis, tuberculosis and many other diseases listed in the specification. The probes and antibodies are also useful for diagnosing cognition and memory disorders, anorexia, hormonal release disorders, cardiovascular activity disorders, pain perception disorders, obesity, diabetes, Alzheimer's disease. Preferably, compounds that decrease or increase the expression of galanin receptor (GAL4) can be used to treat obesity, diabetes, hyperlipidemia and stroke. The GPCR nucleic acid is useful for treating the above mentioned disorders by gene therapy techniques. The present sequence is a novel GPCR polynucleotide of the invention.

Sequence 741 BP; 239 A; 133 C; 158 G; 211 T; 0 other,

#### Alignment Scores:

Pred. No.: 7,06e-135 Length: 741  
Score: 1276.00 Matches: 244  
Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 99.19% Mismatches: 0  
Query Match: 57.09% Indels: 0  
DB: 24 Gaps: 0

us-10-070-241b-1 (1-431) x AAS98050 (1-741)

186 LeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCysCysLeuGluTrp 205  
1 ATGAAGATCAAAATATGACTTCCCTATATGAAAGGAAACACATCTGCTTAGAAGATGG 60  
206 ThrSerProValHisGlnLysIleTyrThrPheIleLeuValIleLeuPheLeuLeu 225  
61 ACCAGTCCGIGGCACACAAACACACACACCTTCATCTCTGCACTTCCTCTCTC 120  
226 ProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrpIleLysLys 245  
121 CCTCTTATGGTGGATGCTTATCTGTAACAGTAAATTTGGTATGAAAGAA 180  
246 ArgValGlyAspGlySerValIleuArgThrIleHisGlyLysGluMetSerLysIleAla 265  
181 AGAGTTGGGAGTGTGCTGCTGCACTATTCATGTAAGAAATTTGCAAAATATGCT 240  
266 ArgLysLysLysArgAlaValIleMetMetValThrValValAlaLeuPheAlaValCys 285  
241 AGGAAGAAGAACAGCTGTCTATATCATGCTGACAGTGGCTGGCTCTCTTTGCTGTGC 300  
286 TrpAlaProPheHisValValHisMetMetIleGluTyrSerAsnPheGluLysGluTyr 305  
301 TGGGACCATTCCTCAATGTGCTCATATGATGATGATGATGATGATGATGATGATGAT 360  
306 AspAspValThrIleLysMetIlePheAlaIleValGlnIleIleGlyPheSerAsnSer 325

361 GATGATGTCACAAATCAAGATGATTTTCTGCTATCGTGCAAAATATTGGATTTTCCAACTCC 420  
326 IleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLysLysAsnValIleuSer 345  
421 ATCTGTAATCCCATTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
346 AlaValCysTyrCysIleValIleAsnLysThrPheSerProAlaGlnArgHisGlyAsnSer 365  
481 GATTTTGTATTTATAGTAATAAATTTCTTCTGAGACAAAGGACATGAGAAATTTCA 540  
366 GlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGluAsnProValGluGlu 385  
541 GGAATTTACAAATCATCGGGAAGAAACAAAGTTTCCCTCAGAGAGAAATCCAGTCGAG 600  
386 ThrLysGlyGluAlaPheSerAspGlyAsnIleGluValLysLeuGlyGluGlnThrGlu 405  
601 ACCAAAGGAGAGACATTCAGTCATGCAACATTTGAAGTCAAAATTTGTCGACACAGAG 660  
406 GluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGluLeuAlaGluAsnSer 425  
661 GAGAAGAAAGAGCTCAAGGACATCTGCTCTCTTTAGGCTGAACTGGCTGAGAAATTTCT 720  
426 ProLeuAspSerGlyHis 431  
721 CCTTTAGACAGTGGGCAT 738  
RESULT 15  
AAD01143  
ID AAD01143 standard; cDNA; 511 BP.  
XX AAD01143;  
AC AAD01143;  
XX 02-NOV-2000 (first entry)  
XX Human orphan G protein-coupled receptor hRUP4 PCR fragment.  
XX Human orphan G protein-coupled receptor hRUP4 PCR fragment.  
KW transmembrane receptor; drug screening; expressed sequence tag; EST;  
KW alternative splicing; PCR fragment; ss.  
OS Homo sapiens.  
XX W0200031258-A2.  
XX 02-JUN-2000.  
XX 13-OCT-1999; 99WO-US23687.  
XX 20-NOV-1998; 98US-0109213.  
XX 16-FEB-1999; 99US-0120416.  
XX 26-FEB-1999; 99US-0121852.  
XX 12-MAR-1999; 99US-0124946.  
XX 12-MAR-1999; 99US-0123949.  
XX 26-MAY-1999; 99US-0136436.  
XX 28-MAY-1999; 99US-0136437.  
XX 28-MAY-1999; 99US-0136439.  
XX 28-MAY-1999; 99US-0136567.  
XX 28-MAY-1999; 99US-0137127.  
XX 28-MAY-1999; 99US-0137131.  
XX 29-JUN-1999; 99US-0141448.  
XX 24-SEP-1999; 99US-0154555.  
XX 29-SEP-1999; 99US-0156633.  
XX 29-SEP-1999; 99US-0156634.  
XX 29-SEP-1999; 99US-0156653.  
XX 01-OCT-1999; 99US-0157280.  
XX 01-OCT-1999; 99US-0157281.  
XX 01-OCT-1999; 99US-0157282.  
XX 01-OCT-1999; 99US-0157293.  
XX 01-OCT-1999; 99US-0157294.  
XX 12-OCT-1999; 99US-0416760.  
XX 12-OCT-1999; 99US-0417044.  
XX

```

Db      |||||
482 AACTCCATCTGTAAATCCCATTTCTCTATGCA 511

Search completed. June 17, 2003, 06.16:31
Job time : 367 secs

```

|    |  |
|----|--|
| PA | (AREN-) ARENA PHARM INC.   |
| XX |  |
| XX | Chen R, Dang HT, Liaw CW, Lin J;   |
| PP |  |
| XX | WPT; 2000 400068/14  |
| XX |  |
| XX | Novel human orphan G protein-coupled receptors and the encoding cDNAs    |
| PT | for use in the identification of G protein-coupled receptor agonists -   |
| PT |  |
| XX | example 1; page 20; 102pp; English.                                      |
| XX |  |
| XX | The present cDNA sequence is the PCR fragment of the endogenous human    |
| CC | orphan G protein coupled receptor (GPCR) hRUP4, obtained from            |
| CC | amplification of human brain cDNA. This fragment is an alternatively     |
| CC | spliced form of the EST (expressed sequence tag) clone A1307658 and is   |
| CC | used to design primers for 5' and 3' RACE of GPCR hRUP4 sequence.        |
| CC | The orphan GPCR of the invention, like all GPCRs has seven transmembrane |
| CC | alpha helices with an extracellular N-terminus and an intracellular      |
| CC | C-terminus. However, no endogenous ligands has yet been identified for   |
| CC | the proteins of the invention. The orphan GPCRs may be used in the       |
| CC | identification of their endogenous ligands, and to screen potential GPCR |
| CC | agonists and antagonists for use as pharmaceutical agents. The proteins  |
| CC | may also be used in the study of GPCR-mediated signalling cascades, and  |
| CC | to elucidate their precise role in normal and diseased human conditions. |
| CC | Nucleic acid encoding human orphan GPCRs may be used for tissue          |
| CC | localisation expression analysis to provide information about their      |
| CC | function in healthy and pathological states.                             |
| CC |  |
| XX | Sequence 511 BP; 143 A; 101 C; 114 G; 153 T; 0 Other;                    |
| XX |  |
| XX |  |

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2003, 05:38:05 ; Search time 1929 Seconds  
(without alignments)  
3618.587 Million cell updates/sec

Title: US 10-070-241B-1  
Perfect score: 2235  
Sequence: 1 MQALNITPEQFSRLPDHNL .....PHLALFRSELAENSPDLSGH 431

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO\_spool/US10070241/runat\_10062003\_095454\_20848/app\_query.fasta\_1.583  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_WAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10070241.acgn.1.1\_1906 -runat\_10062003\_095454\_20848 -NCPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPUP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID        | Description         |
|------------|-------|-------------|--------|----|-----------|---------------------|
| 1          | 1081  | 48.4        | 740    | 13 | B1729969  | B1729969 603550046  |
| 2          | 1006  | 45.0        | 657    | 10 | BB656182  | BB656182 BB656182   |
| 3          | 990   | 44.3        | 636    | 10 | BB663258  | BB663258 BB663258   |
| 4          | 979   | 43.8        | 630    | 10 | BB649191  | BB649191 BB649191   |
| 5          | 902   | 40.4        | 866    | 14 | BQ735771  | BQ735771 AGENCOURT  |
| 6          | 873   | 39.1        | 645    | 10 | BB626475  | BB626475 BB626475   |
| 7          | 782   | 35.0        | 651    | 10 | BB641290  | BB641290 BB641290   |
| 8          | 745.5 | 33.4        | 634    | 10 | BB642180  | BB642180 BB642180   |
| c 9        | 711   | 31.8        | 432    | 9  | AI307658  | AI307658 tb27a03.x  |
| c 10       | 711   | 31.8        | 432    | 9  | AI308124  | AI308124 tb23d09.x  |
| c 11       | 689.5 | 30.9        | 642    | 9  | AL651383  | AL651383 AL651383   |
| 12         | 582   | 26.0        | 730    | 12 | BC169612  | BC169612 602321566  |
| 13         | 558.5 | 25.0        | 672    | 10 | BB632079  | BB632079 BB632079   |
| 14         | 522   | 23.4        | 664    | 10 | BB084541  | BB084541 BB084541   |
| 15         | 460.5 | 20.6        | 3336   | 11 | AK018543  | AK018543 Mus muscu  |
| 16         | 452   | 20.2        | 665    | 10 | BB264499  | BB264499 BB264499   |
| 17         | 450   | 20.1        | 546    | 17 | A2454676  | A2454676 LM0256A18  |
| 18         | 369.5 | 16.5        | 2702   | 11 | AK004730  | AK004730 Mus muscu  |
| 19         | 367   | 16.4        | 646    | 10 | BB277215  | BB277215 BB277215   |
| 20         | 364   | 16.3        | 993    | 14 | BM2626746 | BM2626746 AGENCOURT |
| 21         | 297.5 | 13.3        | 819    | 9  | AU125660  | AU125660 AU125660   |
| 22         | 292   | 13.1        | 2611   | 11 | AK004891  | AK004891 Mus muscu  |
| c 23       | 291   | 13.0        | 525    | 13 | BI133700  | BI133700 UI-M-BH3-  |
| 24         | 289.5 | 13.0        | 1764   | 11 | BC018330  | BC018330 Mus muscu  |
| 25         | 288   | 12.9        | 635    | 14 | BM939496  | BM939496 UI-M-BH3-  |
| 26         | 286   | 12.8        | 599    | 14 | BM933820  | BM933820 UI-M-BH3-  |
| c 27       | 286   | 12.8        | 965    | 17 | CNS0491K  | AL280001 Tetraodon  |
| 28         | 286   | 12.8        | 1070   | 17 | CNS038N1  | AL258772 Tetraodon  |
| 29         | 286   | 12.8        | 1075   | 17 | CNS03800  | AL258777 Tetraodon  |
| 30         | 285   | 12.8        | 637    | 10 | BB632359  | BB632359 BB632359   |
| 31         | 275   | 12.3        | 675    | 9  | AU125231  | AU125231 AU125231   |
| 32         | 274.5 | 12.3        | 2146   | 11 | AK018378  | AK018378 Mus muscu  |
| 33         | 273   | 12.1        | 1017   | 14 | BM22807   | BM22807 AGENCOURT   |
| c 34       | 271   | 12.1        | 1016   | 17 | CNS0407J  | AL302248 Tetraodon  |
| c 35       | 268.5 | 12.0        | 960    | 9  | AL571798  | AL571798 Tetraodon  |
| 36         | 268.5 | 12.0        | 1663   | 11 | BC029291  | BC029291 Homo sapi  |
| 37         | 266   | 11.9        | 935    | 9  | AL531903  | AL531903 AL531903   |
| 38         | 265.5 | 11.9        | 661    | 13 | BM278050  | BM278050 As.tgz.52  |
| 39         | 265   | 11.9        | 3005   | 11 | AK019478  | AK019478 Mus muscu  |
| 40         | 263   | 11.8        | 928    | 9  | AL545857  | AL545857 AL545857   |
| 41         | 262   | 11.7        | 820    | 13 | BI837108  | BI837108 603090071  |
| 42         | 261.5 | 11.7        | 654    | 9  | AU205865  | AU205865 AU205865   |
| c 43       | 260   | 11.6        | 794    | 17 | CNS01WC1  | AL170218 Tetraodon  |
| c 44       | 259   | 11.6        | 2790   | 11 | AK020259  | AK020259 Mus muscu  |
| 45         | 258   | 11.5        | 907    | 9  | AL552677  | AL552677 AL552677   |

ALIGNMENTS

RESULT 1  
B1729969  
LOCUS B1729969 740 bp mRNA linear EST 30-SEP-2001  
DEFINITION 603350046F1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:5357699 5',  
mRNA sequence.  
VERSION B1729969  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 740)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: ecap@rs-remail.nih.gov  
Tissue procurement: The Cepko Laboratory  
cDNA library preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
plate: L1AM11909 row: k column: 12  
high quality sequence stop: 712.  
Location/Qualifiers  
1. 740  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5357699"  
/clone\_lib="NIH\_MGC\_94"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pCMV-SORT6; Site 1: Noti;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: This is a NIH\_MGC Library."  
BASE COUNT 150 a 225 c 184 g 181 t  
ORIGIN

FEATURES  
source

Alignment Scores:  
Pred. No.: 1,39e-122 Length: 740  
Score: 1081.00 Matches: 204  
Percent Similarity: 91.88% Conservative: 11  
Best Local Similarity: 87.18% Mismatches: 19  
Query Match: 48.47% Indels: 0  
Gaps: 0  
US 10-070-241B-1 (1-431) x B1729969 (1-740)

QY 1 MetGlnAlaLeuAsnIleThrProGlnPheSerArgLeuLeuArgAspHisAsnLeu 20  
Db 47 ATGACGGCGCTCAACATCAACACGGCGGACAGTTCCTCCGGCTGAGCGGACACCTG 96  
QY 21 ThrArgGluGlnPheIleAlaLeuValArgLeuArgProLeuValThrProGlnLeu 40  
Db 97 ACTCGGCAATAGCTTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 156  
QY 41 ProGlyArgAlaIleValLeuValLeuValLeuValLeuValLeuValLeuValLeu 60  
Db 157 GCGCGCGCGCTAACTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216  
QY 61 PheGlyAsnAlaLeuValPheValValValValValValValValValValValVal 80  
Db 217 TTGCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276  
QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100  
Db 277 AACATCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336  
QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysIleMet 120  
Db 437 GTACATATCTTCAAT 366  
QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
Db 497 GTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456  
QY 141 ValGluArgIleIleGlyLeuValHisProPheLysMetLysTrpGlnThrTrpAsnArg 160  
Db 457 GTTGAGAGCGGACCAAGGACATCACTCCCTTTTAAATGAAGTCGAGTACATACCCGA 516  
QY 161 ArgAlaPheThrMetLeuGlyValValValValValValValValValValValVal 180  
Db 517 AGGCGCTTCAATATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 576

## QY

181 TrpHisValGlnGlnLeuGlnIleLeuLysTyrAspPheLeuTyrGlnLysGlnHisIleCys 209  
Db 577 TGGCAGGTACAAAGCGTCGAGATTAGTATGACTTCTCTATGAGAAAGACATCTCTCC 636  
QY 201 CysLeuIleGlnTrpThrSerProValHisGlnLysIleTyrThrPheIleLeuVal 220  
Db 637 TGTATGCAAGCTGGAGCAACACCATGAGAAATATATATATATATATATATATATAT 696  
QY 221 LileLeuPheLeuLeuProLeuMetValMetIleLeuLeuTyr 234  
Db 697 ATATCT 738

## RESULT 2

HH656182 657 bp mRNA linear EST 26-OCT 2001  
LOCUS HH656182  
DEFINITION RIKEN full-length enriched, 12 days embryo spinal ganglion  
Mus musculus cDNA clone D130671b2 5', mRNA sequence.  
BB656182  
ACCESSION BB656182 1 GI:16490010  
VERSION  
KEYWORDS  
SOURCE house mouse.  
ORIGIN Mus musculus

## REFERENCE

1 (bases 1 to 657)  
Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,  
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki,  
Y., Okido, T., Saito, P., Sakai, K., Sakai, K., Sano, H., Sasaki,  
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
Tadami, M., Tadawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
Muranatsu, M., and Hayashizaki, Y.  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
Unpublished (2001)

## AUTHORS

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome\_res@sc.riken.go.jp  
URL: <http://genome.asc.riken.go.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subtraction of cDNA libraries for rapid discovery of new  
genes genome Res 10 (10), 1617-1640 (2000)  
Wagil, K., Fujiwara, S., Inoue, K., Toqawa, Y., Izawa, M., Ohara, E.,  
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura,  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.

## TITLE

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer genome Res. 10  
(11), 1757-1771 (2000)

## JOURNAL

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
non-redundant cDNA library genome Res. 11 (2), 281-289 (2001)  
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamazaki, I., Aizawa,  
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.

## COMMENT

Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (<http://genome.asc.riken.go.jp>) for  
further details.  
e mouse tissues.

FEATURES  
source

Location/Qualifiers  
1. 657  
/organism="Mus musculus"  
/db\_xref="taxon:10090"

```

/clone="Dl30067L02"
/clone_lib="RIKEN full-length enriched, 12 days embryo
spinal ganglion"
/tissue_type="spinal ganglion"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGCGCGCAACTCGAGTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGATTCTCGAGTTAATTAAATTCCTCCCGCCGCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from lambda FIC 1."
BASE COUNT      133 a 201 c 159 g 164 t
ORIGIN
Alignment Scores:
Pred. No.:      2 14e-113      Length:      657
Score:          1006.00      Matches:    189
Percent Similarity: 94.04%      Conservative: 16
Best Local Similarity: 86.70%      Mismatches: 13
Query Match:      45.01%      Indels:      0
DB:              10      Gaps:
US-10-070-241B-1 (1-431) x BB656182 (1-657)
QY      24  GlnPheIleAlaLeuValThrArgProLeuValThrProGluLeuProGlyArg 43
Db      1  GAGTTCATCGGCACATATGGGTGGCGGCACATGGTCTACCCCGACGATCTCCCGCGC 60
QY      44  AlaLysLeuAlaLeuValThrGlyValLeuIlePheAlaLeuAlaLeuPheGlyAsn 63
Db      61  GCGAGGCTGGCGCTCTGCTGCTGGCATGCTTCATCTTTGGCCTGGCGCTCTTCGCGAAC 120
QY      64  AlaLeuValPheValThrArgSerLysAlaMetArgThrValThrAsnIlePhe 83
Db      121 GCGCTGGTGTATCTATCTGGTGCACCGCAGCAAGCGCATGCGGCACCTCACCAACATCTTC 180
QY      84  IleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIleProValThrMet 103
Db      181 ATCTGCTCCTGACACTAGCGACCTGCTCATCTCTTCTGCTATCCCGTCCACCATG 240
QY      104 LeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMetValProPhe 123
Db      241 CTCACAGAGCTCTCGACACCTGGCTGGGGGGTGGCTTCATTTGTAAATGCTGCCATTT 300
QY      124 ValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAlaValGluArg 143
Db      301 GTCCAGTGCACCTGCCATCTGACAGAAATCCCTACTATGACCTGACCTGCTGGAAGG 360
QY      144 HisGlnGlyLeuValHisProPheLysMetLysTrpGlnThrThrAsnArgArgAlaPhe 163
Db      361 CACCAGGACTTGTCCATCTTTAAATGAAGAGCGACAGTACCAATCAAGAGCTTTC 420
QY      164 ThrMetLeuGlyValValThrPleuValAlaValIleValGlySerProMetThrHisVal 183
Db      421 ACAATGCTAGTGTGCTGGTGGCTGGTGGCCATTCATCATGAGTATCACCCTGCGCATGTG 480
QY      184 GlnGlnLeuGluIleLysThrAspPheLeuThrGluLysGluHisIleCysCysLeuGlu 203
Db      481 CAGCCACTTCAGATTAACTATGACATCTCCTATATGAAGAAAGAACACATCTGCTGCTGAA 540
QY      204 GluTrpThrSerProValHisGlnIlyIleThrThrPheIleLeuValIleLeuPhe 223
Db      541 GATGTGAGTACAGCTGCTGACAGAGATATACACACTTCATCTCTTCTACCTCTTC 600

```

```

QY      224  LeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeu 241
Db      601  CTGCTACACACAGTTCCTGCTCTCTGTCCTCTACGGCAAAATCGGTATGACGTT 654

RESULT 3
BB663258
LOCUS      BB663258
DEFINITION BB663258 RIKEN full-length enriched, 15 days embryo head Mus
ACCESSION BB663258
VERSION    BB663258.1
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanaoka, T., Hara, A.,
Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, T., Kono, H., Kouda,
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Sasaki,
Okazaki, Y., Okino, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
waqi, K., Fujiwara, S., Inoue, K., Toqawa, Y., Izawa, M., Ohara, E.,
Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real time sequence clustering for construction of a
non-redundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa,
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES
Location/Qualifiers
1..636
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="D930037B12"
/clone_lib="RIKEN full-length enriched, 15 days embryo
head"
/sex="mixed"
/tissue_type="head"
/dev_stage="15 days embryo"
/lab_host="DH10B"

```

/note="Site\_1: Sall; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'

GAGACATAGAAATATGAAAGAGCTTTTITTTTTTITVN 3']. cDNA was prepared by using trachealose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'

GAGACATAGAAATATGAAATATGAAATATGAAATGAGGAGG 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified phlucscript KS(+) after bulk excision from lambda

FLC 1"

BASE COUNT  
ORIGIN

125 a 191 c 149 g 171 t

Alignment Scores:

Pred. No.: 1,93e-111 Length: 636  
Score: 990.00 Matches: 186  
Percent Similarity: 92.89% Conservative: 10  
Best Local Similarity: 88.15% Mismatches: 15  
Query Match: 44.30% Indels: 0  
DB: 10 Gaps: 0

US-10-070-241B 1 (1-411) x HH663258 (1-636)

QY 23 GluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeuProGly 42  
Db 1 GAACAGTTCATCATGCTATGGCTGAGCGCTGCTGCTAATATGAGCTGCGCGG 60  
QY 43 ArgAlaIleLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeuPheGly 62  
Db 61 CGCGCTAACTGTTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 63 AsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThrAsnIle 82  
Db 121 AACTCTCTGGTCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 83 PheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIleProValThr 102  
Db 181 TTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
QY 103 MetLeuGlnAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysLysMetValPro 122  
Db 241 ATGCTCCAGAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
QY 123 PheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAlaValGlu 142  
Db 301 TTCTGCTCAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 143 ArqlHisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArgArgAla 162  
Db 361 AGGACATGAAAGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 163 PheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMetTrpHis 182  
Db 421 TTCACATCTTGGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 183 ValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCysCysLeu 202  
Db 481 GTACACGCTCCAGATTAAGATGACCTTCTCTATGAGAAACACATGCTCTGCTGCTTTC 540  
QY 203 GluGluTrpThrSerProValHisGlnLysIleTyrThrThrPheIleLeuValIleLeu 222  
Db 541 GAAGATGTGGGCGAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 223 PheLeuLeuProLeuMetValMetLeuIleLeu 233  
Db 601 TTCTCTCTGCGCTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633

# RESULT 4

BB649191

LOCUS

DEFINITION

BB649191

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB649191 630 bp mRNA linear EST 26-OCT-2001  
musculus cDNA clone CI30060K24 5', mRNA sequence.

BB649191.1 GI:16483446  
EST.  
house mouse.

Mus musculus  
Fukuyota; Metazoa; Chordata; Craniata; Vertebrata, Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (Bases 1 to 630)  
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanaqaki,T., Hara,A.,  
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

, M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,  
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki

, D., Shibata,K., Shinagawa,A., Shiraki,T., Soqabe,Y., Suzuki,H.,  
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,  
Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
Unpublished (2001)

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@qsc.riken.go.jp,  
URL: http://genome.qsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

, M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagl,K., Fujisake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura

, S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

, Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a

non-redundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa

, K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
Hayashizaki,Y.  
Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.qsc.riken.go.jp>) for  
further details.  
e mouse tissues.

Location/Qualifiers  
1..630  
/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="c130060K24"

/clone\_lib="RIKEN full length enriched, 16 days embryo  
head"

/sex="mixed"

/tissue\_type="head"

/dev\_stage="16 days embryo"

/lab\_host="DH10B"

/note="Site\_1: Sall; Site\_2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken

## FEATURES

source



contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGGATCCCAAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCGAGTTTAAATTAATTCCTCCGCCCCGCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmidscript KS(+) after bulk excision from Lambda FLC 1".

BASE COUNT 128 a 193 c 151 g 158 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4, 35e-110 Length: 630  
Score: 979.00 Matches: 183  
Percent Similarity: 94.74% Conservative: 15  
Best Local Similarity: 87.56% Mismatches: 11  
Query Match: 43.80% Indels: 0  
DB: 10 Gaps: 0

US-10-070-241B-1 (1-431) x BB649191 (1-630)

QY 33 ProLeuValTyrThrProGluLeuProGlyArgAlaLysLeuAlaLeuValLeuThrGly 52  
DB 4 CCACCTGGTGCCTCACCCCGCAGCTTCCCGCGCGCGCAGCTGGCCCTCCTGCTGCTGGC 63

QY 53 ValLeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValValThrArg 72  
DB 64 ATGCTCATCTTTGGCCTGGCGCTCTTCGGCAACGCCCTGGTAGCTATGTGGTGACCGGC 123

QY 73 SerLysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeu 92  
DB 124 AGCAAGGCATCGCGACCGTCACCAACATCTTCATCTGCTCCCTGGCACTCAGCGACTG 183

QY 93 LeuIleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAspAsnTrpLeu 112  
DB 184 CTCATCTGCTTCTTCTGTCATCCGCTCCACATGCTCCAGAACGCTCCGACACCTGGCTG 243

QY 113 GlyGlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGlu 132  
DB 244 GGGGTGCTTCATTTGCAAAATGGTCCATTTGTCCAGTGCATGCTATTGTGACAGAA 303

QY 133 IleLeuThrMetThrCysIleAlaValGluArgHisGlnGlyLeuValHisProPheLys 152  
DB 304 ATCCTTACTATACCTGCTGCTGTGGAAGGACCAAGGACCTGTGCTCCTTTTAAA 363

QY 153 MetLysTrpGlnTyrThrAsnArgAlaPheThrMetLeuGlyValValTrpLeuVal 172  
DB 364 ATGAAGCGCAGCTACACCAATCAAGAGCTTTCACAATGCTAGGTGGTGGTGGTGGT 423

QY 173 AlaValIleValGlySerProMetTrpHisValGlnGlnLeuGluIleLysTyrAspPhe 192  
DB 424 GCCATCATCATAGGATCACCATGTCGATGTCGAGGACCTTGAGATTAGTATGACTTC 483

QY 193 LeuTyrGluLysGluHisIleCysCysLeuGluGluTrpThrSerProValHisGlnLys 212  
DB 484 CTATATGAAAAGAACACATCTGCTGCTGGAAGAGTGGAGCAGCCCGCTGCACCAAG 543

QY 213 IleTyrThrThrPheIleLeuValIleLeuPheLeuLeuProLeuMetValMetLeuIle 232  
DB 544 ATCTACACCACTTCATCTTGTCCACCTCTTTCCTGCTTACCACGTGCTGCTCTCTGTC 603

QY 233 LeuTyrSerLysIleGlyTyrGluLeu 241  
DB 604 CTCTACGGGAAATCGGTTATGAGCTT 630

RESULT 5

BQ735771

LOCUS

DEFINITION AGNCN0PT\_8064527 NICHD XG5 Emb4 Xenopus laevis cDNA clone IMAGE:5571086 5', mRNA sequence.

BQ735771  
BQ735771.1 GI:21874668  
EST.  
SOURCE African clawed frog.  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.  
1 (bases 1 to 866)  
NCI-CoAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapps@email.nih.gov  
Tissue Procurement: Dr. Igor Dawid  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: NCI-CoAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Plate: LLAM12314 row: f column: 15  
High quality sequence stop: 553.  
Location/Qualifiers  
1..866  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone="IMAGE:5571086"  
/dev\_host="embryo, stage 31-32"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1; NotI: Site 2; SalI: Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

BASE COUNT 218 a 216 c 195 g 237 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2, 29e-100 Length: 866  
Score: 902.00 Matches: 183  
Percent Similarity: 82.93% Conservative: 21  
Best Local Similarity: 74.39% Mismatches: 28  
Query Match: 40.36% Indels: 15  
DB: 14 Gaps: 3

US-10-070-241B-1 (1-431) x BQ735771 (1-866)

QY 30 ArgLeuArgProLeuValTyrThrProGluLeuProGlyArgAlaLysLeu-AlaLeuVa 49  
DB 28 CGGTCGCGCCACGCGTC-----CGCCACGCGCTCCGCTTGT 66

QY 49 lLeuThrGlyValLeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrVa 69  
DB 67 GGCAATCTGCTGCTCATATTTCTGTCGCGCTTGTGGCAATCTCTCTACTTATGT 126

QY 69 lValThrArgSerLysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLe 89  
DB 127 GGTGACCAAGAACCAATGAGGACTGTCACCAACATCTTCATCTGCTCTGGCGCT 186

QY 89 uSerAspLeuLeuIleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAs 109  
DB 187 CAGTGATCTACTCATTTGGCTTCTTCGATCCCTTCAGCATGCTGCACCAACATCTCC 246

QY 109 pAsnTrpLeuGlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaVa 129  
DB 247 CAAGTGGCTGGAGGAGGGTTCGCTTGCATAATGGTTCCCTTGTTCAGCTCTATGCC 306

QY 129 lValThrGluIleLeuThrMetThrCysIleAlaValGluArgHisGlnClyLeuValHi 149  
DB 307 TGTACTGAAATCTGACATGACATGTCATTTGCTGTGGAAGGCATCAAGGATGTACA 366



Db 291 AACATCTTCATCTGCTCTGGCACTCAGTGATCTGCTCATTTGGCTTCTTCTGCAATCCC 350  
 QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysIysMet 120  
 Db 351 GTACAGATGCTCCAAACATCTCCACAAAGTGGCTGGTGGCTTCATCTGCAAGATG 410  
 QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140  
 Db 411 GTGCGCTTCGTCACCTCCACAGCTGCTGTGTGACGGAAATCCCTACCATGCTGATCCCT 470  
 QY 141 ValGluArgHisGlnGlyLeuValHisProPheIysMetIysTrpGlnTrpThrAsnArg 160  
 Db 471 GTTGAGAGGCACCAAGGATCTCATCCATCTTTTAAATGAACTGCTAGTACATACCTGA 530  
 QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180  
 Db 531 AGGGCTTTTACAAATCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 590  
 QY 181 TrpHisValGlnGlnLeuGluIleLysTrpAspPheLeuTrpGluLysGlu 197  
 Db 591 TGGCAGCTACAAAGCTTCGAGATTAG-TATGACTTCTCTATGAGAAAGAC 640

## RESULT 7

BB641290 651 bp mRNA linear EST 26-OCT-2001  
 BB641290 RIKEN full-length enriched, 10 days neonate cortex Mus  
 musculus cDNA clone A830015G12 5', mRNA sequence.

BB641290.1 GI:16476415

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
 Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Konda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,  
 Okazaki, Y., Okido, T., Saito, P., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
 Muramatsu, M., and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel.: 81-45-503-9222

Fax: 81-45-503-9216

E-mail: genome-res@gsc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

M., Konno, H., Okazaki, Y., Muramatsu, M., Sugahara, Y., Shibata, K., Itoh

Carninci, P., Shibata, Y., Hayashizaki, Y., and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura

S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and

Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara

Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa

K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and

Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

## FEATURES

source

Location/Qualifiers

1..651

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="A830015G12"

/clone\_lib="RIKEN full-length enriched, 10 days neonate

cortex"

/tissue\_type="cortex"

/dev\_stage="10 days neonate"

/lab\_host="DH10B"

/note="Site 1. Sall, Site 2. BamHI, cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN, Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. cDNA went through one round of normalization

to Rot = 20.0 and subtraction to Rot = 459.0. Second

strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGATCTCGATTAAATAATTAATCCCCCCCCCCC

3']. cDNA was cleaved with XhoI and BamHI. Vector: a

modified pBluescript KS(+) after bulk excision from Lambda

FLC I."

BASE COUNT

131 a

193 c

151 g

172 t

4 others

ORIGIN

Alignment Scores:

Pred No :

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-070-241B-1 (1-431) x BB641290 (1-651)

QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgGluArgPheLeuValThrPheGluLeu 40

Db 4 ACTGGGAAAGTTCATTATCGTATGAGGTTAGAGAGTGTGTCTATACATTTGGAGCTG 63

QY 41 PheGlyArgAlaIleCysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeu 60

Db 64 CCCCGCGCGCTAAACTGGCTTTGGCTGGCTGGAGCACTCATTTTTCCTGGCGCTC 123

QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80

Db 124 ITTGGTAACCTTGGTCACTATGTTATGTTGAGTATGTTTATTGCTTTTATCTGACAC 183

QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspIleGluIleThrPhePheCysIlePro 100

Db 184 AACATCTTCATCTGCT 243

QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysIysMet 120

Db 244 GTACAGATGCTCCAGAACATCTCCGACCAAGTGGCTGGTGGCTTCATCTGCAAGATG 303

QY 121 ValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAla 140

Db 304 GTGAGATTCGTCAGTCT 363

QY 141 ValGluArgHisGlnGlyLeuValHisProPheIysMetIysTrpGlnTrpThrAsnArg 160

Db 364 GTTGAGAGGCACCAAGGATCTCATCCATCTTTTAAATGAACTGCTAGTACATACCTGA 423



```

QY 161 ArqAlaPheThrMetLeuGlyValValTrp-LeuValAlaValIle-ValGlySerPro 180
Db 528 CGCGCTTACAAATCTGGGTGCTGCTGGTGGCAACCATCGGAGGATCACCCC 587
QY 180 eTrpHisValGln-GlnLeuGluIleLysTyrrAspPheLeuVal 194
Db 588 TGGGGGAGGTACAAAGGCTTGAATAAAGTTGATTCCTCTAT 632

RESULT 9
A1307658/c 432 bp mRNA linear EST 08-APR-1999
LOCUS tb27a03.x1 NCI-CCAP_Kid12 Homo sapiens cDNA clone IMAGE:2055185 3'
DEFINITION similar to SW:GALR_RAT Q62805 GALANIN RECEPTOR ;, mRNA sequence.
ACCESSION A1307658
VERSION A1307658.1 GI:4001837
KEYWORDS EST.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi,
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 432)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CCAP clone distribution information can be
found through the I M A G E Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1414 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 419.
Location/Qualifiers
1. 432
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2055185"
/clone_lib="NCI-CCAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/notes="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker, Site 1: Not 1, Site 2: Eco RI,
Plasmid DNA from the normalized library NCI-CCAP_Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 129 a 92 c 92 g 119 t
ORIGIN
Alignment Scores:
Pred No. Length 432
Score: 711.00 Matches: 143
Percent Similarity: 78.57% Conservative: 0
Best Local Similarity: 78.57% Mismatches: 0
Query Match: 31.81% Indels: 39
DB. 9 Gaps: 2

US-10-070-241B-1 (1-431) x A1307658 (1-432)

QY 160 ArgArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerPro 179
Db 432 CGAAGGGCTTTCACAATGCTAGGTGCTGGTGGTGGGATCATGCTAGGATCACCCC 373

```

```

QY 180 MetTrpHisValGlnGlnLeuGluIleLysTyrrAspPheLeuValGluLysGluHisIle 199
Db 372 ATGCGGACGCTGGCAACATTGAGATCAAAATATGACTTCTCTATATCAAAAGGAACACATC 313
QY 200 CysCysLeuGluCysTrpThrSerProValHisClnLysIleTyrrThrThrMetLeu 219
Db 312 TGTGTATTAGAAATGTGACCAAGAGCTGTGTAATAAATATTAATACCTTCACCTT 253
QY 220 ValIleLeuThrLeuLeuProLeuMetValMetLeuIleLeuTyrrSerLysIleClyTyrr 239
Db 252 CTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 232
QY 240 GluLeu-TrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGly 259
Db 231 ---CTTATGC----- 225
QY 259 sCLeuMetSerLysIleAlaArgLysLysIleArgAlaValIleMetMetValThrVal 279
Db 224 -----AAGAAGAAACAGCTGTCATTATGATGATGATGATGATGATGATGATGATGAT 187
QY 279 ValAlaLeuPheAlaValCysTrpAlaTrpPheHisValValHisMetMetIleClyTyrrSe 299
Db 186 GGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 127
QY 299 rAsnPheGluLysGluTyrrAspValThrIleLysMetIlePheAlaIleValGlnI 319
Db 126 TAAATTTTGAAGAGAAATATCATGATGTCACAAATCAAGATGATTTTTCCTATCGTGCAAT 67
QY 319 eIleGlyPheSerAsnSerIleCysAsnProIleValTyrrAlaPheMetAsnGluAsnPh 339
Db 66 TATTGGATTTCCTAAATCTATCTTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7
QY 339 eLys 340
Db 6 CAAA 3

RESULT 10
A1308124/c 432 bp mRNA linear EST 08-APR-1999
LOCUS tb23d04.x1 NCI-CCAP_Kid12 Homo sapiens cDNA clone IMAGE:2055185 3'
DEFINITION similar to SW:GALR_RAT Q62805 GALANIN RECEPTOR ;, mRNA sequence.
ACCESSION A1308124
VERSION A1308124.1 GI:4002759
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 432)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CCAP clone distribution information can be
found through the I M A G E Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1308 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 429.
Location/Qualifiers
1. 432
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2055185"
/clone_lib="NCI-CCAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"

```

/lab\_host="DHI10B"  
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with  
 a modified polylinker, Site 1: Not I, Site 2: Eco RI,  
 plasmid DNA from the normalized library NCL\_CGAP\_Kid5 was  
 prepared, and ss circles were made in vitro. Following RAP  
 purification, this DNA was used as tracer in a subtractive  
 hybridization reaction. The driver was PCR-amplified cDNAs  
 from a pool of 5,000 clones made from the same library  
 (clones 132912-1325831, 1471368-1472903 and  
 1492104-1493255). Subtraction by Bento Soares and M.  
 Fatima Honaldo."

BASE COUNT 129 a 92 c 92 q 119 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,19c-77 Length: 432  
 Score: 711.00 Matches: 143  
 Percent Similarity: 78.57% Conservative: 0  
 Best Local Similarity: 78.57% Mismatches: 0  
 Query Match: 41.81% Indels: 39  
 Gaps: 2

US-10-070-241b-1 (1-431) x A1308124 (1-432)

QY 160 ArqArqAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerPro 179  
 DB 412 CGAAGGCTTTTACAAATGTTAGCTGTGCTGTGGCTGGCAGCATCGTAGGATCAACC 373  
 QY 180 MetTrpHisValGlnGlnLeuGlnIleTyrAspPheLeuTyrGluTyrSerLysHisIle 199  
 DB 372 ATGTGGCAGCTGCAACATTTGAGATCAAAATATATATCTTCTATATGAAAGCAACATC 313  
 QY 200 CysCysLeuGlnGlnTrpThrSerProValHisGlnIleTyrThrThrPheIleLeu 219  
 DB 312 TCGCTGCTTAGAAGACTGGACAGCCCTGTGCACCAAGCATCTACACCTTCATCCTT 253  
 QY 220 ValIleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyr 239  
 DB 252 GTCACTCTCTTCCCTCCCTCCCT----- 232  
 QY 240 GluLeu-TrpIleTyrLysArgValGlyAspGlySerValLeuArgThrIleHisGlyIle 259  
 DB 241 -CTTATGG----- 225  
 QY 259 sCiuMetSerLysIleAlaArgLysLysLysArgValIleMetMetValThrValVa 279  
 DB 224 -----AACAAGAAACGAGCTGTCTCATATATGATGGTGCACAGTGGT 187  
 QY 279 lAlaLeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSe 299  
 DB 186 GGCCTCTCTTCTGT 127  
 QY 299 rAsnPheGluTyrGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnI 319  
 DB 126 TAAATTTGAAAGCAAT 67  
 QY 319 eIleGlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnph 339  
 DB 66 TATTGGATTITTCACCTCATCTGTAATCCCATATGCTATGCAATTTATGAAGAAACCTT 7  
 QY 339 eLys 340  
 DB 6 CAAA 3

## RESULT 11

AL651383

LOCUS

DEFINITION AL651383 XGC-gastrula Silurana tropicalis cDNA clone Tgas040116 5',

mRNA sequence.

## ACCESSION

AL651383

VERSION AL651383.1 GI:17661440

KEYWORDS EST.

SOURCE western clawed frog

## ORGANISM

Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 Xenopodinae; Silurana.

## REFERENCE

1 (bases 1 to 642)

Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (10\_2001)

Unpublished (2001)

## JOURNAL

Contact: Huckle E

## COMMENT

Sanger Centre

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS\_SEQUENCE\_ID: Tgas040116.sp6

Sequencing primer: SP6

## FEATURES

source

1..642

Location/Qualifiers

/organism="Silurana tropicalis"

/db\_xref="taxon:8364"

/clone="Tgas040116"

/clone\_lib="XGC gastrula"

/dev\_stage="gastrula (stages 10.5-13 mixed)"

/lab\_host="Escherichia coli XL1-blue"

/note="Vector: pCS107; Site\_1: EcoRI; Site\_2: NotI; cDNA

was oligo dt primed from 5' of poly A. RNA from stages

10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated

into pCS107 with EcoRI at the 5' end and NotI at the 3'

end."

BASE COUNT 193 a 126 c 131 q 191 t 1 others

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,62c-74 Length: 642  
 Score: 689.50 Matches: 138  
 Percent Similarity: 83.51% Conservative: 19  
 Best Local Similarity: 73.40% Mismatches: 28  
 Query Match: 30.85% Indels: 4  
 Gaps: 1

US-10-070-241b-1 (1-431) x AL651383 (1-642)

QY 238 GlyTyrGluLeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHis 257  
 DB 16 GCGTATCAGCTTTGGATTAA-AAGCGTGTGGTGTGCTTCACTACTGCAACTATTCTAT 74  
 QY 258 GlyLysGluMetSerLysIleAlaArgLysLysLysArgAlaValIleMetMetValThr 277  
 DB 75 GGAAGTCAATGCTCTAAATAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 134  
 QY 278 ValValAlaIlePheAlaValCysTrpAlaProPheHisValValHisMetMetIleGlu 297  
 DB 135 GTGTTGCTCTTTTCTGCACTCTGCTGGGCACTTTCATGTTGTTTCATATGATAGAA 194  
 QY 298 TyrSerAsnPheGluLysGluTyrAspValThrIleLysMetIlePheAlaIleVal 317  
 DB 195 TACAGTAACCTTTGCAAAATCAATATGATGATGATGATGATGATGATGATGATGATGATG 254  
 QY 318 GlnIleIleGlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGlu 337  
 DB 255 CAAATAATGGATTTTCACTCCATTTCCACCCCAATAGTTTATGCTTTTATGAAATGA 314  
 QY 338 AsnPheLysLysAsnValIleuSerAlaValCysTyrCysIleValAsnLysThrPheSer 357  
 DB 315 AACTTCAAGCAAGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374  
 QY 358 ProAlaGlnArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSer 377  
 DB 375 CCTCCAGAGAGCCAGCAAACTCTGCATCACTGATCACTGATCACTGATCACTGATCACT 434  
 QY 378 LeuArgGluAsnProValGluGluThrLysGlyGluAlaPheSerAspGlyAsnIleGlu 397



Mapping of 19042 mouse cDNAs on mouse chromosomes. J. Struct. Funct. Genomics 2 pre, 172-186 (2001). Please visit our web site (<http://genome.qsc.riken.go.jp>) for further details.

```

FEATURES             source
location/Qualifiers
1..672
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A230966L21"
/clone_lib="RIKEN full-length enriched, adult male
hypothalamus"
/sex="male"
/tissue_type="hypothalamus"
/dev_stage="adult"
/lab_host="D110B"
/notes="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GACACACAGACAGATCTCGAGAGCTCTTTTTTTTTTTTTTTTNN 3'], cDNA was
prepared by using triethose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap trapper. cDNA went through one round of normalization
to Rot - 20.0 and subtraction to Rot - 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GACACAGAGATCTCGAGATTAATTAAATTCCTCCCTCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from lambda
FLC I."
115 a 217 c 186 g 146 t 8 others

```

| Alignment Scores:      |          |               |     |
|------------------------|----------|---------------|-----|
| Prod. No.:             | 4, 30-58 | Length:       | 672 |
| Score:                 | 558.50   | Matches:      | 122 |
| Percent Similarity:    | 62.64%   | Conservative: | 9   |
| Best Local Similarity: | 58.47%   | Mismatches:   | 25  |
| Query Match:           | 24.99%   | Indels:       | 54  |
| OH:                    | 10       | Gaps:         | 1   |

US 10 070 241B-1 (1 431) x BB632079 (1-672)

[illegible]

25 PheTleAlaLeuTyrArqLeuArqProLeuValTyrThrProGluLeuProGlyArqAla 44

299 TTTATCCGCAATATATGAGTGGGATATGATGCTTACCTCTGAGTGGGCTGC 311

45 LysIeuAlaIeuValIeuThrGlyValIlePheAlaIeuAlaIeuPheGlyAsnIa 64  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

319 AGGTGGCCCTCTGCTGGTGGATGATCATCTTTGAGATGAGAGTCTTTGGGAAACGGC 378

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

82

379 CTGGTACTGATCTGTGATCCCGCAGGATCATGGCCAGCTACCAACATCTTCATC 438

85 CysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIleproValThrMetLeu 104

439 TGTGCTGGCACTCAGGAGCTGGTCATCGTCTTCTCTGCATCCGCGTCACCATGCTC 498

**Cy**    105 GlnAsnTrlGcTASPASnrrPleuGlYAlaphelIeCysLysMetVa|ProPhcVal | 124  
    III : :: :: :: :: :: :: :: :: ::  
**Cyb**    499 CAGAAANGTCGGGCACACTGGCTGGGG- ---- - 525

125 GluSerThrAlaValValThrGluLeuLeuThrMetThrCysIleAlaValGluArgHis 144  
525

|           |            |   |                 |
|-----------|------------|---|-----------------|
| 145       | QY         | ClnGlyIeuValHisProPheLysMetLysTTrpGlnTyrThrAsnArqArgAlaPheThr   | 164             |
| 525       | Db         | -----   | 525             |
| 165       | QY         | MetLeuGlyValValTTrpLeuValAlaValIlleValGlySerProMetTrpHisValGln  | 184             |
| 526       | Db         | -----GGTGTGGTGTCGTTGGCCATCATNATAGGATCACCCATGTGGCATGTGCAC  | 579             |
| 185       | QY         | GlnLeuGlnIleLysTyrAspPheLeuTyrGlnLysGlnHisIleLysCysLeuGlnGlu  | 204             |
| 580       | Db         | GGACTTCGAGTTAAGTATGATTCTTA-TATTAATAAAGACACATTGGTGGCTGGGAACAG  | 638             |
| 205       | QY         | TrpThrSerProValHisGlnLysIleC  | 213             |
| 639       | Db         | TGGGGTACGCGCGGGAACCAAAAAATTT  | 665             |
| RESULT 14 |            |   |                 |
| HH084541  | LOCUS      | 664 bp  | linear          |
| HH084541  | DEFINITION | RIKEN full-length enriched, adult male diencephalon Mus musculus cDNA clone 9330190620 3', mRNA sequence. | EST 31-AUG-2000 |

BB084541  
ACCESSION  
H084541.2 GI:15407665  
VERSION  
EST.  
KEYWORDS  
house mouse.  
SOURCE  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata, Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 664)  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanaoka, K., Hara, A.

Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasakih, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Arawaka, T., et al. 2001)  
 TITLE  
 JOURNAL  
 COMMENT  
 This is sequence version released 2001-06-07  
 on Jun 21, 2000

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration, Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
http://genome.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000).

waiki, K., Fujiwaki, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

K., Kamakura, S., Kiyosawa, H., Kohno, S., Saito, F., Shinada, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y.  
Mapping of 19032 mouse cDNAs on mouse chromosomes J. Struct. Func. Genomics 2 pre, L72-L86 (2001)  
Please visit our web site (<http://genome.ysc.riken.go.jp/>) for further details.



cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group, in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

# FEATURES

```

Location/Qualifiers
1..664
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="RIKEN full-length enriched, adult male
diencephalon"
/sex="male"
/tissue_type="diencephalon"
/dev_stager="adult"
/lab_host="DH10H"
/Note="Site 1: SalI, Site 2: BamHI, cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCCAAGAGCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot -10.0 and subtraction to Rot -185.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATCTCGAGTAAATTAATTAATCGGCGGCGGCGG
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified phagescript KS(+) after bulk excision
from Lambda FIC I. Cloning sites, 5' end: SalI; 3' end:
BamHI"
BASE COUNT      216 a 110 c 135 g 202 t 1 others
ORIGIN
Alignment Scores:
Pred. No.:
Score: 1.36e-53 Length: 664
Percent Similarity: 81.51% Matches: 104
Best Local Similarity: 71.23% Conservative: 15
Query Match: 23.14% Mismatches: 27
DB: 10 Indels: 2
Gaps: 0
US-10-070-241B-1 (1-431) x BH084541 (1-664)
QY 286 TrpAlaProPheHisValValHisMetMetIleGluTyrSerAsnPhcGluLysGluTyr 305
Db 2 TGGGCACCTTCNCACTGCTTTTCACATGATGTTGCTGAGTACAGTAACCTT-GAAAAAGAGTAT 59
QY 306 AspAspValThrIleLysMetIlePheAlaIleValGlnIleIleGlyPheSerAsnSer 325
Db 60 GATGATGTCACATCAAGATGGTTTCTGCTGTTCACACAAACAAATGGCTTTTCACTCC 119
QY 326 IleCysAsnProIleValTyrAlaPheMetAsnGluAsnPhcLysLysAsnValSer 345
Db 120 ATCTGTAATCCCTTTGCTGTATGCATTTATGAATGAAACTTCAAAAGAAATTTTGTCT 179
QY 346 AlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArgHisGlyAsnSer 365
Db 180 GCGGTTTGTATTGCACTAGCAAGAACCTTCTCCCGAGGACACACCTCGGAAATTCCT 239
QY 366 GlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGluAsnProValGluGlu 385
Db 240 GGGATTTCAATGATCAAAAGAGAGCAAGATTTATCAGCATCAGCGTCCAGTGGCGAA 299
QY 386 ThrLysGluAlaPheSerAspGlyAsnIleGluValLysLysCysGluClnThrGlu 405
Db 300 GCCAAGAGGACATTATCAGCGATGCCAATGCCAGGTGATGTCAATATGTGTGAGCAGCCAGG 359
QY 406 GluLysLysLysLysLeuLysArgHisIleLeuAlaLeuPheArgSerGluLeuAlaGluAsnSer 425

```

```

16 360 GAGAAAAAGGAAATTAAGGACAGATTTGAGTTTATATTAATTTTATATAAAATTT 419
QY 426 ProLeuAspSerGlyHis 431
Db 420 ACTTTCGGCAGTGGAGAT 437
RESULT 15
AK018543 3336 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus adult male colon cDNA, RIKEN full-length enriched
DEFINITION library, clone:9030607L22:neuropeptide Y receptor Y2, full insert
sequence.
ACCESSION AK018543
VERSION AK018543.1 GI:12858296
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male colon cDNA to mRNA,
clone:9030607L22:neuropeptide Y receptor Y2, full insert
sequence.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1
Garninci,P and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
942794/53
10349636
2
Garninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojohori,T., Bono,H., Kasikawa,T., Saito,P.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,
Quackenbush,J., Schriml,M., Stauble,F., Suzuki,P., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsch,G., Blake,T., Boiffelli,D., Bolund,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.T., Bull,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Humphreys,A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombarts,P., Nordone,P.,
Ring,P., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyooka,K., Wanq,K.H., Weitz,C., Whitaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawai,J., Kohtsuki,S.
and Hayashizaki,Y.
Functional annotation of a full length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

```



```
Db      1341 CCTGGAA-----GTCAAAAGAGACAATGGCCGACACTGACTCTTTTCGGAGGTAC 1391
QY      395 -AsnIleGluValLysLeuCysGluGlnThr 404
Db      1392 CAATGTGTAAAGCACACAGGTGTGAAGCACA 1422
```

Search completed: June 17, 2003, 06:49:03  
Job time : 1956 secs





DB 106 CTGACTCAGATTAATTAATAC-----TATGCAACTACTACTTCCACGAGCT 156  
QY 34 LeuValTyrThrProGluLeuProGlyArqAlaLysLeuAlaLysLeuValLeuThrGlyVal 53  
DB 157 CAAGTG-----GCAGCAATCTTCATTATTCTCTACTTT 189  
QY 54 LeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValValThrArgSer 73  
DB 190 CTGATCTCTTTTCTGATATATGAGGAATA-TGFGGTTTGCTTTATTGTAATAGGAAC 249  
QY 74 LysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeu 93  
DB 250 AAACATATGCCACACTACTAACTCTTCATCTTAAACCTGCCCATAGCTACTTACTA 309  
QY 94 IleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGly 113  
DB 310 GTTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369  
QY 114 GlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGluIle 133  
DB 370 GAAACACGATGTGAAATAGTGTATGTGAGAAATATGTTGAAATATGTTGAAATATGTT 424  
QY 144 LeuThrMetThrCysIleAlaValGluArqIleGlnGlyLeuValIleHisProPheLysMet 153  
DB 430 TTTATGTTAGTTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489  
QY 154 LysTrpGlnTyrThrAsnArqAlaPheThrMetLeuGlyValValThrLeuValAla 173  
DB 490 AAG-----CTCATATCATGAGACAGGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
QY 174 ValIleValGlySerPro-----MetTrpHisValGlnGlnLeuGluIleLysTrp 190  
DB 544 ATCAGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 191 AspPheLeuTyrGluLysLysIle-----IleCysCysLeuLeuGluTrp 205  
DB 601 CCACTGACACTCAACTCCACAGATAAAACCAGCTCCAGCTACTGCTGCCGCGGAGACTGG 660  
QY 206 ThrSerProValHisGlnLysIleTyrThrPheIleLeuValIleLeuPheLeuLeu 225  
DB 661 CCAATACGAAATGAGAAATATACAAATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720  
QY 226 ProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrpIleLysLys 245  
DB 721 CCGCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771  
QY 246 ArqValGlyAspGlySerValLeuArqThrIleHisGlyLysGluMetSerLys----- 263  
DB 772 -----AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816  
QY 264 -----IleAlaArqLysLysArqAlaValIleMetMetMetValThrValValAlaLeu 281  
DB 817 CAGCTGGTGTCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 876  
QY 282 PheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsnPhe 301  
DB 877 TTTATCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 946  
QY 302 GluLysGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnIleIleLys 321  
DB 937 TCTCAATGAACTGCCATCATCAACATCATCAACATCATCAACATCATCAACATCATCAAC 996  
QY 322 PheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLysLys 341  
DB 997 TTGCGCAACACATGCTCAATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1056  
QY 342 AsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArg 361  
DB 1057 GGTTCGAAGAAGCTTTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116  
QY 362 HisGlyAsnSerGlyIleThrMetMetMetMetMetMetMetMetMetMetMetMetMet 381  
DB 1117 TATGCC-----CTAAAGCTAAAGCCCATGCTGCTCATAAAC 1152

QY 382 ProValGluGluThrLysGlyGluAla---PheSerAspGlyAsnIleGluValLysLeu 400  
DB 1153 ACATCTAATCAGCTTGCCAGGAATCATCATTTTCAAAACCTCATGAGGAACCTTGCTT 1212  
QY 401 CysGluGlnThrGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420  
DB 1213 TATAGGAAAGTGTGTAAGAAACCCCAACAGGAA-----TTAGTGTGGAAGAA 1260  
QY 421 LeuAlaGluAsnSer 425  
DB 1261 TTAAGAAACTACT 1275  
RESULT 2  
US-08-687-355A-3  
Sequence 3, Application US/08687355A  
Patent No. 5989834  
GENERAL INFORMATION:  
APPLICANT: Synaptic Pharmaceutical Corporation  
TITLE OF INVENTION: NUCLEIC ACID ENCODING NPHROPPPTIDE  
TITLE OF INVENTION: V/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,355A  
FILING DATE: No. 5989834ember 26, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 44742-A-PCF/JHW/MAT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1556 base pairs  
TYPE: nucleic acid  
STRANDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 211..1353  
US-08-687-355A-3  
Alignment Scores:  
Pred. No.: 2,32e-48 Length: 1556  
Score: 486.00 Matches: 126  
Percent Similarity: 49.52% Conservative: 81  
Best Local Similarity: 30.14% Mismatches: 143  
Query Match: 21.74% Indels: 68  
Gaps: 15  
US-10-070-241b-1 (1-431) x US-08-687-355A-3 (1-1556)  
QY 17 AspHisAsnLeuThrArqGluGlnPheIleAlaLeuTyrArqLeuArqProLeuValTyr 36  
DB 235 GATGCAATCAAACTCTAGAACTGAAAGTGGAACTCTATGCTGGGGGCCCACTCTCT 294

QY 37 ThrProGluLeuProGly-----Arg 43  
 Db 295 AGAGGTGAGTTCGCCCTGATCCAGCGGAGCTCATAGACAGACCAAACTGGTTGAG 354  
 QY 44 AlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeuPheGlyAsn 63  
 Db 355 GTGCAGGTGGTCTTATACCTGCTTATGTTCCATCATCTTCTGGCGGTAGTTGGCAAC 414  
 QY 64 AlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThrAsnIlePhe 83  
 Db 415 TCTCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474  
 QY 84 IleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIleProValThrMet 103  
 Db 475 ATTGCCAACCTGGCTGGCGGATCTTTTGGTGAACACCTGTGCTGCCATTCACCTT 534  
 QY 104 LeuGlnAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysLysMetValProPhe 123  
 Db 535 ACCTATACCTTGTATGGGGAGTGGAATGGTCCAGTTTCTGCCATTTGGTGGCCUAT 594  
 QY 124 ValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAlaValGluArg 143  
 Db 595 GCCCAGGTCTGGCAGTCAAGTGTCCACATACTTTGACAGTCATTCCTTTGGACCGA 654  
 QY 144 HisGlnGlyLeuValHisProPheLysMetLysTrpGlnIleThrAsnArgAlaPhe 163  
 Db 655 CATCGTTGCATTTGTACCACTGGAGCAAG-----ATCTCCAAAGCAAAATCAGCTTC 708  
 QY 164 ThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMetTrpHisVal 183  
 Db 709 CTGATTATTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 768  
 QY 184 GlnGln-----LeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 Db 769 CGGAGTACTCACTGATTGAGATTATTCCTGCATTT-----GAGATTGTAGCC 816  
 QY 201 CysLeuGluGluTrp-----ThrSerProValHisGlnLysIleTyrThrThrPhe 217  
 Db 817 TGTACTGAGAAATGACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 876  
 QY 218 IleLeuValIleLeuPheLeuProLeuMetValMetLeuIleLeuTyrSerLysIle 237  
 Db 877 ACCCTGCTAAATCCCTAGCTTTTCTGCTGGCATCATATCTTCTCTACACCGGATC 936  
 QY 238 GlyTyrGluLeuTrp-----IleLysLysArgValGlyAspGlySerValLeuArgThr 255  
 Db 937 -----TGGAGTAAGCTAAAGAAACCAAGCTAGTCTGTGAGGTGCAAGTGACCAT 984  
 QY 256 IleHisGlyLysGluMetSerLysIleAlaArgLysLysLysArgAlaValIleMetMet 275  
 Db 985 TACCAT-----CAGCGAAGGCACAAACAGCAGCAAAATGCTC 1020  
 QY 276 ValThrValValAlaLeuPheAlaValCysTrpAlaProPheHisValValHisMetMet 295  
 Db 1021 GTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080  
 QY 296 IleGluTyrSerAsnPheGlu-----LysGluTyrAspValThrIleLysMet 312  
 Db 1081 GTGACATCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1125  
 QY 313 IlePheAlaIleValGlnIleIleGlyPheSerAsnSerIleCysAsnProIleValThr 332  
 Db 1126 ATCTTACCGGTGTCACATATTGCGATGTGCTCCACCTTTCGCCCAACCCCTCTCTAT 1185  
 QY 333 AlaPheMetAsnGluAsnPheLysLysAsnValLeuSerAlaValCysTyrCysIleVal 352  
 Db 1186 GCGTGGATCAACAGCACTACAGAAAGCTTTCTCTCTCAGCC----- 1227  
 QY 353 AsnLysThrPheSerProAlaGlnArg-----HisGlyAsnSerGlyIleThr 368  
 Db 1228 -----TTCGGCTGTGAGCAGAGGTGGATGCCATTCACCTCGAGAGGTGTCCATGACC 1278  
 QY 369 MetMetArgLysLysAlaLysPheSerLeuArgGluAsnProValIleGluThrLysGly 388

Db 1279 TTC-----AAGCTAAAGAACACCTGGCA-----GTCAAAAGAACAAATGCG 1320  
 QY 389 -----GluAlaPheSerAspGlyAsnIleGluValLysLeuCysGluGlnThr 404  
 Db 1321 CTCACAGACTCTTTTTCAGAGCGCCACCAACGCGTGAAGAAAGCTGCTCAAGAGAG 1374  
 RESULT 3  
 US-09-407-367-3  
 : Sequence 3, Application US/09407367  
 : Patent No. 6420532  
 : GENERAL INFORMATION:  
 : APPLICANT: Christophe P.G. Gerald, et al.  
 : TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2 SPECIFIC COM  
 : NUMBER OF SEQUENCES: 27  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cooper & Dunham LLP  
 : STREET: 1185 Avenue of the Americas  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: U.S.A.  
 : ZIP: 10036  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/407,367  
 : FILING DATE:  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: White, John P.  
 : REGISTRATION NUMBER: 28,678  
 : REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPW  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 212-278-0400  
 : TELEFAX: 212-391-0525  
 : INFORMATION FOR SEQ ID NO: 3:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1556 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: DNA (genomic)  
 : HYPOTHEICAL: NO  
 : ANTI-SENSE: NO  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 211..1353  
 : US-09-407-367-3  
 Alignment Scores:  
 Pred. No.: 2,326-48 Length: 1556  
 Score: 486.00 Matches: 126  
 Percent Similarity: 49.52% Conservative: 81  
 Best Local Similarity: 30.14% Mismatches: 143  
 Query Match: 21.74% Indels: 68  
 DB: 4 Gaps: 15  
 US-10-070-241B-1 (1-431) x US-09-407-367-3 (1-1556)  
 QY 17 AsnHisAsnLeuThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyr 36  
 Db 235 GATGAGATCAAACTGTAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 294  
 QY 37 ThrProGluLeuProGly-----Arg 43  
 Db 295 AGAGGTGAGTTCGCCCTGATCCAGCGGAGCTCATAGACAGACCAAACTGGTTGAG 354  
 QY 44 AlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeuPheGlyAsn 63  
 Db 355 GTGCAGGTGGTCTTATACCTGCTTATGTTCCATCATCTTCTGGCGGTAGTTGGCAAC 414





```

Db      319  ATTGCCAACCTGGCTGGTGGGATCTTTTGGTCAACACCTGGTGGCTGCCTACCTCTT 378
QY      104  LeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMetValProPhe 123
Db      379  ACCATATCCTTGGTGGGAGTGGAAATGGTCCAGTTTGTGCACTTGGTGGCTAT 438
QY      124  ValGlnSerThrAlaValThrGluIleLeuThrMetThrCysIleAlaValGluArg 143
Db      439  GCGCAGGCTGGCAGTCAAGTGTCCACAACTATCTTACACGTCATCTTGGACCA 409
QY      144  HisGlnGlyLeuValHisProPheLysMetLysTrpGlnTrpThrAsnArgArgAlaPhe 163
Db      499  CATCGTTGCATGTCTACCACTGGAGCAAG-----ATCTCCCAAGCAAAATCAGCTTC 552
QY      164  ThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMetIrpHisVal 183
Db      553  CTGATTAATGGCTGGCTGGGTCGAGCGCTCTGCTGCAAGTCCCTTGGCATCTTC 612
QY      184  GlnGln-----LeuGluIleLysTrpAspPheLeuTrpGluLysGluHisIleCys 200
Db      613  CGGGAGTACTCACTGATTGAGATTATTCCTGACTT-----GAGATTGTAGCC 660
QY      201  CysLeuGluGluTrp-----ThrSerProValHisGlnLysIleTrpThrPhe 217
Db      661  TGTACTGAGAATGGCCCGGGAGGAGAGAGTGTGTACGGTACACTACAGCCCTTCC 720
QY      218  IleLeuValIleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTrpSerLysIle 237
Db      721  ACCCTGTAACTCTACCTTTGCTTGGCTTGGCTGGCTATATCTTCTTCCACCCCGATC 780
QY      238  GlyTrpGluLeuTrp-----IleLysLysArgValGlyAspGlySerValLeuArgThr 255
Db      781  -----TGGAGTAAGCTAAAGAACACAGCTTAGTCTCGAGCTGCAAGTGACCAT 828
QY      256  IleHisGlyLysGluMetSerLysIleAlaArgLysLysLysArgAlaValIleMetMet 275
Db      829  TACCAT-----CAACGAGGCGACAAATGACCAAAATGCTC 864
QY      276  ValThrValValAlaLeuPheAlaValCysTrpAlaProPheHisValValHisMetMet 295
Db      865  CTGTCGCTGGTGGTGGTGTGGCAGTCAGCTGGCTGCCCTCCATGCCCTCCAACTTGC 924
QY      296  IleGluTrpSerAsnPheGlu-----LysGluTrpAspValThrIleLysMet 312
Db      925  GTGACATGACAGCACTACAGAAAGCTTCTCTCTCAGCC-----AAACTC 369
QY      313  IlePheAlaIleValGlnIleIleGlyPheSerAsnSerIleCysAsnProIleValTrp 332
Db      970  ATCTTCACCGTGTCCACATATTGGCATGTGTGTCACCTTCCCAACCCCTCTCTAT 1029
QY      333  AlaPheMetAsnGluAsnPheLysLysAsnValLeuSerAlaValCysTrpCysIleVal 352
Db      1030  GGCTGGATGAACCACTACAGAAAGCTTCTCTCTCAGCC----- 1071
QY      353  AsnLysThrPheSerProAlaGluArg-----HisGlyAsnSerGlyIleIleHis 368
Db      1072  -----TTCCGCTGTGACAGAGTTGGATTTGCATTAATCTGTAAGTGTGATSAAG 1122
QY      369  MetMetArgLysLysAlaLysPheSerLeuArgGluAsnProValGluGluThrLysGly 388
Db      1123  TTCAAGGCTAAAGAACCTGGAAGTCAAAAGAACAAATGGCTCACTGACTCTTTTCA 1182
QY      389  GluAla 390
Db      1183  CAGGCC 1188

```

## RESULT 5

US-09-407-367-5

; Sequence 5, Application US/09407367

; Patent No. 6420532

; GENERAL INFORMATION:

; APPLICANT: Christophe P.G. Gerald, et al.

```

; TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2 SPECIFIC COM
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,367
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/POCKET NUMBER: 44742-AA-PCT-US/JPW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOL. CULF. TYPE: genomic DNA
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 55..1200
; US-09-407-367-5

```

## Alignment Scores:

```

Pred No : 4 25e-47      Length: 1200
Score: 474.00      Matches: 119
Percent Similarity: 49.00%      Conservative: 78
Best local Similarity: 29.60%      Mismatches: 145
Query Match: 21.21%      Indels: 60
DB: 4      Gaps: 12

```

US-10-070-241b-1 (1-431) x US-09-407-367-5 (1-1200)

```

QY      17  AsphHisLeuThrArgGluGlnPheIleAlaLeuTyArgLeuArgProLeuValTyr 36
Db      79  GATGAAATCAAAATATAGAAAGTGAAGTGAATTTATATGAGTGGAGCAACATCTCT 138
QY      37  ThrProGluLeuProGly-----Arg 43
Db      139  AGAGCTGAGTTCGCCCTGATGCCAGACCCGAGCTCAACACAGCACCAAACTGGTTGAG 198
QY      44  AlaLysLeuAlaLeuValLeuIleGlyValLeuIlePheAlaLeuAlaLeuPheGlyAsn 63
Db      199  GTCCAGGTGGTCCCTATCTACTGGCCTATGTTCCATCATCTGCTGGCGCTAGTTCGCAAC 258
QY      64  AlaLeuValPheTyValValThrArgSerLysAlaMetArgThrValThrAsnIlePhe 83
Db      259  TCTCTGTTAATCCATGTGTGATCAAAATTCAGAGCATGCGACAGTAAACCACTTTT 318
QY      84  IleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIleProValThrMet 103
Db      319  ATTGCCAACCTGGCTGGGAGTCTTTTGGTGAACACCCCTGGCTGCTCACTCTCT 378
QY      104  LeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMetValProPhe 123
Db      379  ACCATATCCTTGGTGGGAGTGGAAATGGTCCAGTTTGTGCACTTGGTGGCTAT 438

```

QY 124 ValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAlaValGluArg 143  
 DB 439 GNCAGGGCTGACAGTAAATGTCACAAATAATTTGCAAGTATTTGATTTGAAACA 498  
 QY 144 HisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArqAlaPhe 163  
 DB 499 CATCGTCGTCGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 552  
 QY 164 ThrMetLeuGlyValValThrLeuValAlaValIleValGlySerProMetTrpHisVal 183  
 DB 553 CTGATTATTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 612  
 QY 184 GlnGln-----LeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200  
 DB 613 GGGAGTATATCTATTGATATTATTTATTT-----GAGATTATAGAT 660  
 QY 201 CysLeuGluGluTrp-----ThrSerProValHisGlnLysIleTyrThrThrPhe 217  
 DB 661 TGTATGTGACAAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 720  
 QY 218 IleLeuValIleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIle 237  
 DB 721 ACCCTGCTANTCTCTAGCTTTTCTCTCTGAGATCATATTTTCTCTATACAGCGATC 780  
 QY 238 GlyTyrGluLeuTrp-----IleLysLysArgValGlyAspGlySerValLeuArgThr 255  
 DB 781 -----TGGATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 828  
 QY 256 IleHisGlyLysGluMetSerLysIleAlaArqLysLysLysArqAlaValIleMetMet 275  
 DB 829 TACCAT-----CAGCGAAGGCACAAATGACCAAAATGCTC 864  
 QY 276 ValThrValValAlaLeuPheAlaValCysTrpAlaProPheHisValValHisMetMet 295  
 DB 865 GT 924  
 QY 296 IleGluTyrSerAspPheGlu-----LysGluTyrAspAspValThrIleLysMet 312  
 DB 925 GTGGACATCGACAGATATGTTGTGATCTGAAGGAGTAT-----AAATC 969  
 QY 313 IlePheAlaIleValGlnIleIleGlyPheSerAsnSerIleCysAsnProIleValTyr 332  
 DB 970 ATCTTCACCGCTTCCACATATATATGATGATGATGATGATGATGATGATGATGAT 1029  
 QY 333 AlaPheMetAsnGluAsnPheLysLysAsnValLeuSerAlaValCysTyrCysIleVal 352  
 DB 1040 GATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071  
 QY 353 AsnLysThrPheSerProAlaGlnArg-----HisGlyAsnSerGlyIleThr 368  
 DB 1072 -----TTCTGTCTCTCAATATATATATATATATATATATATATATATATAT 1122  
 QY 369 MetMetArqLysLysAlaLysPheSerLeuArqGlnAsnProValGluGluThrLysGly 388  
 DB 1123 TTCAAGGCTAAAAACAATGTCGATCAATCAATCAATCAATCAATCAATCAATCAAT 1182  
 QY 389 GluAla 390  
 DB 1183 GAGGCC 1188

## RESULT 6

US 08-192-288-1

Sequence 1, Application US/08192288

Patent No. 554549

GENERAL INFORMATION:

APPLICANT: Gerald, Christophe

APPLICANT: Walker, Mary

APPLICANT: Branche, Theresa

APPLICANT: Weisshank, Richard L.

TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE

TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTOR AND USES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 2

: CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cooper & Dunham  
 : STREET: 40 Rockefeller Plaza  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: U.S.A.  
 : ZIP: 10112  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/192,288  
 : FILING DATE:  
 : CLASSIFICATION: 514  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: White, John P.  
 : REGISTRATION NUMBER: 28,678  
 : REFERENCE/DOCKET NUMBER: 44742/JFW/TEP  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (212) 977-9550  
 : TELEFAX: (212) 664-0525  
 : TELEX: 422523 COOP UI  
 : INFORMATION FOR SEQ ID NO: 1:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1280 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 43..1185  
 : US-08-192-288-1

## Alignment Scores:

Pred. No.: 1,62e-46 Length: 1280  
 Score: 469.50 Matches: 113  
 Percent Similarity: 51.65% Conservative: 75  
 Best Local Similarity: 31.04% Mismatches: 123  
 Query Match: 21.01% Indels: 53  
 DB: 1 Gaps: 13

US-10-070-241B-1 (1-431) x US-08-192-288-1 (1-1280)

QY 38 ProGlnLeuProGlyArqAlaLysLeu-----AlaLeuValLeuThrGly 52  
 DB 154 CCAGCTTATACATAGTACCAAGCTGATGAGGATACAACTGCTGTCATATTGCGCTAC 213  
 QY 53 ValLeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValValThrArg 72  
 DB 214 TGATCATCATCATCTGTTGAGTTAATTTGAAATTTGATTTGATTTGATTTGATTTGAT 273  
 QY 73 SerLysAlaMetArgThrValThrAsnIlePheIleCysSerIleAlaLeuSerAspLys 92  
 DB 274 TTCAGAGATATGCGACATACCAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTT 333  
 QY 93 LeuIleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAsnTrpLeu 112  
 DB 334 TTGCTGAACACATCTGTCTCTACCGTTCATCTTACCTATACCTTAATAGGCGGAGTGGAA 193  
 QY 113 GlyGlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGlu 132  
 DB 394 ATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453  
 QY 133 IleLeuThrMetThrCysIleAlaValGluArqHisGlnGlyLeuValHisProPheLys 152  
 DB 454 ACAAATACCTTAAATTAATTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 513  
 QY 153 MetLysTrpGlnTyrThrAsnArqAlaPheThrMetLeuGlyValValThrLeuVal 172  
 DB 514 AGCAAG-----ATCTCCAGAGCAATACGCTTCTGATTATTTGGCTTGGCTGGGCGCATC 567

```

QY 173 AlaValIleValGlySerProMetThrPheValGlnGln-----LeuGluIleLys 189
Db 568 AGTGGCCCTGCTGGCAAGTCTCTGTCATCTTCTGAGAGATATTGCTATTGAGATATC 627
QY 190 TyrAspPheLeuTyrGluLysGluHisIleCysCysLeuGluGluTrp-----Thr 206
Db 628 CCGGACTTT-----GAGATTGCGCTGCTACTGAAATCTGCTGAGGAGGAG 675
QY 207 SerProValHisGlnLysIleTyrThrPheIleLeuValIleLeuPheLeuPro 226
Db 676 AAGACATCTATGCGCACTGCTATAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCT 735
QY 227 LeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrp-----IleLys 244
Db 736 CTGGGCATTATATCATTTTCTACACTCGCAT-----TGGAGTAAATGGAAG 783
QY 245 LysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGluMetSerLysIle 264
Db 784 AACCATGTCTAGCTCTGAGCTGCAAAATGACCACTACCAT-----822
QY 265 AlaArgLysLysLysArgAlaValIleMetMetValThrValValAlaLeuPheAlaVal 284
Db 823 ---CAGCAAGCAAAACCAACCAATGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 879
QY 285 CysTrpAlaPropPheHisValValHisMetMetIleGluTyrSerAsnPheGlu-----302
Db 880 AGCTGGTCTGCTCTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
QY 303 ---LysGluTyrAspValThrIleLysMetIlePheAlaIleValIleIleGly 321
Db 940 CTGAAGGAGTAC-----AACTCATCTTCACAGTGTTCACATCATCGCC 984
QY 322 PheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPheLysLys 341
Db 985 ATGTCTCCACTTTTCCCAATCCCTTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
QY 342 AsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArg 361
Db 1045 GCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
QY 362 -----HisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSer 377
Db 1078 TTGAGAGCATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
QY 378 LeuArgGluAsn 381
Db 1135 GTGAGAGAGAAC 1146

```

## RESULT 7

```

US-08-687-355A-1
; Sequence 1, Application US/08687355A
; Patent No. 5989834
; GENERAL INFORMATION:
; APPLICANT: Syntactic Pharmaceutical Corporation
; TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE
; TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,355A
; FILING DATE: No. 5989834,ember 26, 1996

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPW/MAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 278-0400
; TELEFAX: 212 391-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..1185
; US-08-687-355A-1

```

## Alignment Scores:

```

Pred. No.: 1,62e-46 Length: 1280
Score: 469.50 Matches: 113
Percent Similarity: 51.65% Conservative: 75
Best Local Similarity: 31.04% Mismatches: 123
Query Match: 21.01% Indels: 53
DB: 2 Gaps: 13

```

US-10-070-241B-1 (1-431) x US-08-687-355A-1 (1-1280)

```

QY 38 ProGluLeuProGlyArgAlaLysLeu-----AlaLeuValLeuThrGly 52
Db 154 CCAGAGCTTATAGATAGTACCAAGCTGATTGAGGTACAAAGTGTCTCATATTGGCCTAC 213
QY 53 ValLeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValValThrArg 72
Db 214 TGCCTCATCATCTTGGTGGGTAAATGGCAACTCTCTGTCATCATCTGTTGATCAAA 273
QY 73 SerLysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeu 92
Db 274 TTAAAGCATGCGCAAGTAAACCAATTTTTCATTATTAATCTGCTGTGACAGATCTT 333
QY 93 LeuIleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAsnTrpLeu 112
Db 334 TTTTAAACATCTGTCTCTACCTTCTTAACTTAACTTAACTTAACTTAACTTAACTTAA 393
QY 113 GlyGlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGlu 132
Db 394 ATGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 453
QY 133 IleLeuThrMetThrCysIleAlaValGluArgHisGlnGlyLeuValHisProPheLys 152
Db 454 ACAATTAATTTGAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
QY 153 MetLysTrpGlnTyrThrAsnArgArgAlaPheThrMetLeuGlyValValTrpLeuVal 172
Db 514 AGTAAG-----ATCTCAAGAGCAATCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 567
QY 173 AlaValIleValGlySerProMetThrPheHisValGlnGln-----LeuGluIleLys 189
Db 568 AGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 627
QY 190 TyrAspPheLeuTyrGluLysGluHisIleCysCysLeuGluGluTrp-----Thr 206
Db 628 CCGGACTTT-----GAGATTGCGCTGCTACTGAAATCTGCTGAGGAGGAG 675
QY 207 SerProValHisGlnLysIleTyrThrPheIleLeuValIleLeuPheLeuPro 226
Db 676 AAGACATCTATGCGCACTGCTATAGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 735
QY 227 LeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrp-----IleLys 244
Db 736 CTGGGCATTATATCATTTTCTACACTCGCAT-----TGGAGTAAATGGAAG 783

```



QY 322 PheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPhelLysLys 341  
 Db 985 ATGTCTCCACTTTTGGCAATCCCTCTCTATGCTGGATGAACACCACTACAGAAAG 1044  
 QY 342 AsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArg 361  
 Db 1045 GCTTCTCTCTCGGC-----TTCCGCTGTGACGAGGG 1077  
 QY 362 -----HisGlyAsnSerClyIleThrMetMetArgLysLysAlaLysPheSer 377  
 Db 1078 TTGGATGCCATTCATCTGAGCTGCTGACATTCAGAGCTAAAGG-----AACCTCCAG 1134  
 QY 378 LeuArgGluAsn 381  
 Db 1135 GTCAGAAAGAAC 1146

## RESULT 9

US-08-876-798A-1  
 : Sequence 1, Application US/08876798A  
 : Patent No. 6355478  
 : GENERAL INFORMATION:

: APPLICANT: Baez, Melvyn  
 : APPLICANT: Yang, Peiyi  
 : TITLE OF INVENTION: RHESUS MONKEY NEUROPEPTIDE Y RECEPTOR  
 : NUMBER OF SEQUENCES: 3  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Eli Lilly and Company  
 : STREET: Lilly Corporate Center  
 : CITY: Indianapolis  
 : STATE: Indiana  
 : COUNTRY: U.S.A.  
 : ZIP: 46285

: COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent In Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/876,798A  
 : FILING DATE: 16-JUN-1997  
 : CLASSIFICATION: 435

: ATTORNEY/AGENT INFORMATION:  
 : NAME: Gaylo, Paul J.  
 : REGISTRATION NUMBER: 36,808  
 : REFERENCE/DOCKET NUMBER: X-10901  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (317)276-0756  
 : TELEFAX: (317)276-2763

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2144 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 63..1205  
 : US-08-876-798A-1

Alignment Scores:  
 Pred. No.: 4,82e-46 Length: 2144  
 Score: 468.50 Matches: 112  
 Percent Similarity: 51.65% Conservative: 76  
 Best Local Similarity: 30.77% Mismatches: 123  
 Query Match: 20.96% Indels: 53  
 DB: 4 Gaps: 13

US-10-070-241B-1 (1-431) x US-08-876-798A-1 (1-2144)

QY 38 ProGluLeuProGlyArgAlaLysLeu-----AlaLeuValLeuThrGly 52  
 Db 174 CCAGAGCTATAGATAGTACCAAGCTGATTGAGTACAAAGTTGTCTCATATTGGGCTAT 233

QY 53 ValLeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValValThrArg 72  
 Db 234 TGCTCCATCATCTTGGTGGGTAATTTGCCAACTCTGTGATCCACCTGGTGAICAAA 293  
 QY 73 SerLysAlaMetArgThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeu 92  
 Db 294 TTCAGAGCATGGCAGTAAACCACTTTTATGCCAACTGGCTGTGGCAGATCTT 353  
 QY 93 LeuIleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAspAsnTrpLeu 112  
 Db 354 GTGTGAATACTCTGTGTCTACCATCTCTTACCTTACCTTAAATGGGGAGTGGAAA 413  
 QY 113 GlyGlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGlu 132  
 Db 414 ATGGGTCTCTCTGTGGCCACCTGGTGGCCCTATGCCACAGGGCTGCCAGTATCC 473  
 QY 133 IleLeuThrMetThrCysIleAlaValGluArgHisGlnGlyLeuValHisProPheLys 152  
 Db 474 ACAATCACTTGTACAGTAATTTGGCTTGAAGGATAGGTGTATGATACACCTGGAG 533  
 QY 153 MetLysTrpGlnTyrThrAsnArgAlaPheThrMetLeuGlyValValTrpLeuVal 172  
 Db 534 AGCAAG-----ATCTCAAGCGTATCAAGCTTTTGTATTATTAGCTTGGATGGGCATC 587  
 QY 173 AlaValIleValGlySerProMetTrpHisValGlnGln-----LeuGluIleLys 189  
 Db 588 AGTGGCTCTGTACCAAGTCCCTGGCCATCTCCGGGAGTATTCACCTGATGAGATCAT 647  
 QY 190 TyrAspPheLeuTyrGluLysGluHisIleCysCysLeuGluGluTrp-----Thr 206  
 Db 648 CGGSAITTT-----GAGATTTGTGCTGTATTAATAAATGAGTGTGGAGGAA 695  
 QY 207 SerProValHisGlnLysIleTyrThrPheIleLeuValIleLeuPheLeuPro 226  
 Db 696 AAGAGCATCTATGGCACTGTCTACAGTCTTCTCTCTGTGATCTGTACCTTTTGCT 755  
 QY 227 LeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrp-----IleLys 244  
 Db 756 CTGGGCATAATATCATTTTCTACCTCCCAT-----TGGACTAAATTTGAAG 803  
 QY 245 LysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGluMetSerLysIle 264  
 Db 804 AGCATGTCTAGTCTCGAGCTGCAAAATGACCACTACCAT-----842  
 QY 265 AlaArgLysLysLysArgAlaValIleMetMetValThrValAlaAlaLeuPheAlaVal 284  
 Db 843 ---CAGCGAAGGCAAAAACCCACCAAAATGCTGTGTGGTGTGTGTGTGTGTGTGTGTGT 899  
 QY 285 CysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsnPheGlu-----302  
 Db 900 AGCTGGCTGCTCTCCATGCTTCCAGCTTGGGTTGACATTCACAGCCATGTCCTGGAC 959  
 QY 303 ---LysGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnIleGly 321  
 Db 960 CTGAAGAGTAGTAC-----AAATATATTTCAAGTGTGTGTGTGTGTGTGTGTGTGT 1004  
 QY 322 PheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPhelLysLys 341  
 Db 1005 ATGTGCTCCACTTTTGGCAATCCCTCTCTATGGCTGGATGAACAGTAACTATAGAAAG 1064  
 QY 342 AsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGlnArg 361  
 Db 1065 GCTTCTCTCTCTGCC-----TTCCGCTGTGACGAGGG 1097  
 QY 362 -----HisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSer 377  
 Db 1098 TTGATGCTATTCTACTCTGAGGTGTCTGTCACATTCACAGGCTAAAGG---AACCTGGAG 1154  
 QY 378 LeuArgGluAsn 381  
 Db 1155 GTCAGAAAGAAAT 1166



; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA: US/09/119,788  
 ; FILING DATE: 21-JUL-1998  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/053,790  
 ; FILING DATE: 25-JUL-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: King, William T  
 ; REGISTRATION NUMBER: 30,954  
 ; REFERENCE/DOCKET NUMBER: GH50029  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-270-5515  
 ; TELEFAX: 610-270-5090  
 ; TELEX:

INFORMATION FOR SEQ ID NO. 1.

SEQUENCE CHARACTERISTICS:

LENGTH: 1633 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

US-09-119-788-1

Alignment Scores:

|                        |          |               |      |
|------------------------|----------|---------------|------|
| Pred No :              | 2 85e-45 | Length:       | 1633 |
| Score:                 | 460.50   | Matches:      | 110  |
| Percent Similarity:    | 52.31%   | Conservative: | 94   |
| Best Local Similarity: | 28.21%   | Mismatches:   | 146  |
| Query Match:           | 20.60%   | Indels:       | 40   |
| DB:                    | 4        | Gaps:         | 9    |

US-10-070-241b-1 (1-431) x US-09-119-788-1 (1-1633)

|    |     |  |     |
|----|-----|--|-----|
| Qy | 4   | LeuAsnIleThrProGluInPheSerArgLeuLeuArgAspHisAsnLeuThrArgGlu  | 23  |
| Db | 174 | CTGAGTAACTCAAGAGCCCTTT-----TTAAACCCCGCCGACTATGACGACGAG       | 224 |
| Qy | 24  | GlnPheIleAlaLeuValThrArgProLeuValThrProGluLeuProGlyArg       | 43  |
| Db | 225 | GAATTCCTGGG---TACCTGTGGAGGAACTACTGCAC-----CCGAAGAA           | 269 |
| Qy | 44  | AlaLeuValPheValThrArgSerLysAlaMetArgThrValThrAsnIlePhe       | 83  |
| Db | 270 | TATGAGTGGCTCTGATCGCGGCTACATCATCTGCTGCTGCTCTCAATTGGCAAC       | 329 |
| Qy | 64  | AlaLeuValPheValThrArgSerLysAlaMetArgThrValThrAsnIlePhe       | 83  |
| Db | 330 | GTCCTGTTTGTGGCAGTGTGGAGACACACACATGAGGACGCTAACCACTACTTC       | 389 |
| Qy | 84  | IleCysSerLeuAlaLeuSerAspLeuLeuIleThrPheCysIleProValThrMet    | 103 |
| Db | 390 | ATAGTCAATCTTCTGTGGTATGCTGTGGACCATCACTGCCTTCGAGCCACACTG       | 449 |
| Qy | 104 | LeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMetValProPhe | 123 |
| Db | 450 | GTCCTGATATCACTGACACTGCTTTTGGACAGTCCCTTTGCCAAAGTCATTCTCTAT    | 509 |
| Qy | 124 | ValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAlaValGluArg | 143 |
| Db | 510 | CTACAGACCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG      | 569 |
| Qy | 144 | HisGlnGlyLeuValHisProPheLysMetIleTrpGlnThrAsnArgAlaPhe       | 163 |
| Db | 570 | TGGTATGCAATCTGACCTCTTTGATGTTTAAAG-----AGCACAGCAAAAGGCGCGCT   | 623 |
| Qy | 164 | ThrMetLeuGlyValValThrLeuValAlaValIleValGlySerProMetTrpHisVal | 183 |
| Db | 624 | AACAGCAATGTCATCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG     | 683 |
| Qy | 184 | GlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIle-----CysCys  | 201 |

|    |      |  |      |
|----|------|--|------|
| Db | 684  | ATGGAGTCTAGTACCGGTGTCACAGACTTATAGCAATAAATCAATCTTTTACGAGTGT   | 743  |
| Qy | 202  | IeuGluGluTrpThrSerProValHisGlnLysIleThrThrPheIleLeuValIle    | 221  |
| Db | 744  | GATGAGCGCTGGGTGGTAAATTTATCCCAAGATGTACACATCTGTTCTTTCTGTGTG    | 803  |
| Qy | 222  | LeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeu | 241  |
| Db | 804  | ACATACAGCCACCACTGCTCATGTGTGCTTATCTCCAAATATTTCCGAAATCTC       | 863  |
| Qy | 242  | TrpIleLysArgValGlyAspGlySerValLeuLeuArgThr-----              | 255  |
| Db | 864  | TGGTGTGACACATCCCTGGACATCATCTGTAGTTCAGAGNAATGGAAGCCCTGCAG     | 923  |
| Qy | 256  | -----IleHisGly   | 258  |
| Db | 924  | CCGCTTTTACAGCTCTGAGGCTATACAGAGCAATGAAATGCGGATGAGGCTGTGGG     | 983  |
| Qy | 259  | LysGluMetSerLysIleAlaArgLysLysLysArgAlaValIleMetLeuValThrVal | 278  |
| Db | 984  | GCTGAAATTAAGCAGATCCGACCCAGAGAAAGACCCGCGATGCTGATGCTT---GTG    | 1040 |
| Qy | 279  | ValAlaLeuPheAlaValCysTrpAlaProPheHisValHisMetIleGlu---       | 297  |
| Db | 1041 | CTTTTGGTATTGCAATTTGCTATCTACCAATTAGCTCTCAATGCTGCTAAAGAGAGTA   | 1100 |
| Qy | 298  | TyrSerAsnPheGluLysGluTyrAspAspValThrIleLysMetIlePheAlaIleVal | 317  |
| Db | 1101 | TTTGGGATTTTGGCCCATACTGAAAGACAGAGATGTTGTTATTTTATTTTTCAT       | 1160 |
| Qy | 318  | GlnIleIleGlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsn-Gl | 337  |
| Db | 1161 | CACCTGCTTGTATATCCCAATAGCTGCGCAATCAATTTATTTTATTTTCTCAGTCGA    | 1220 |
| Qy | 337  | uAsnPheLysLysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSe | 357  |
| Db | 1221 | AAATTCGAGAGGAATTTA-----AAGCTGGTTTCTGCTGTGCTTGGATTCAG         | 1274 |
| Qy | 357  | rProAlaGlnArgHisGlyAsnSerGly                                 | 366  |
| Db | 1275 | CATCGCCAGAGGATCGGCTCACCAGG                                   | 1302 |

RESULT 12

US-09-255-368-1  
 ; Sequence 1: Application US/09-255-368  
 ; Patent No. 6262246  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gerald, Christophe P.G.  
 ; APPLICANT: Jones, Kenneth A.  
 ; APPLICANT: Honini, James A.  
 ; APPLICANT: Borowsky, Beth  
 ; TITLE OF INVENTION: DNA Encoding Mammalian Neuronal Receptors  
 ; FILE REFERENCE: 1795/57155-A  
 ; CURRENT APPLICATION NUMBER: US/09/255,368  
 ; EARLIER FILING DATE: 1999-02-22  
 ; EARLIER APPLICATION NUMBER: 09/161,113  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SOFTWARE: Patent Ver. 2.0 - beta  
 ; SEQ ID NO 1  
 ; LENGTH: 1410  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 US-09-255-368-1

Alignment Scores:  
 Pred No : 1 48e-42 Length: 1410  
 Score: 437.00 Matches: 101  
 Percent Similarity: 47.50% Conservative: 70  
 Best Local Similarity: 28.06% Mismatches: 119  
 Query Match: 19.55% Indels: 70

```

Db: 1048 GCCAACACCATCATCATAGCTACAGAGAAATTGCGGAGAGATTCCAGAGTACC 1107
US-09-255-368-7
: Sequence 7, Application US/09255368
: Patent No. 6252246
: GENERAL INFORMATION:
: APPLICANT: Gerald, Christophe P.G.
: APPLICANT: Jones, Kenneth A.
: APPLICANT: Bonini, James A.
: APPLICANT: Borowsky, Beth
: TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors
: TITLE OF INVENTION: and Uses Thereof
: FILE REFERENCE: 1795/57155-A
: CURRENT FILING DATE: 1999-02-22
: EARLIER APPLICATION NUMBER: 09/161,113
: EARLIER FILING DATE: 1998-09-25
: NUMBER OF SEQ ID NOS: 42
: SOFTWARE: Patent In Ver. 2.0 - beta
: SEQ ID NO 7
: LENGTH: 1293
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-255-368-7

Alignment Scores:
Pred. No.: 4,476-42 Length: 1293
Score: 432.50 Matches: 93
Percent Similarity: 50.32% Conservatives: 64
Best Local Similarity: 29.81% Mismatches: 126
Query Match: 19.35% Indels: 29
DB: 4 Caps: 5

US-10-070-241b-1 (1-431) x US-09-255-368-7 (1-1293)

QY 49 ValLeuThrGlyValLeuLeuPheAlaLeuPheGlyAsnAlaLeuValPheTyr 68
Db 139 ATGTGGCCATATGCCCTCATCTTCCTGCTCATGCTGGCCACACCTGCTGTTTC 198
QY 69 ValValThrArgSerLysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAla 88
Db 199 ATCGTGCTCAAGAACGGGACATCATATATATATATATATATATATATATATAT 258
QY 89 LeuSerAspLeuLeuLeuThrPhePheCysIleProValThrMetLeuGlnAsnIleSer 108
Db 259 GTCAATGACCTGTGTGTGATATCTTGTGATATATATATATATATATATATATAT 318
QY 109 AspAsnTrpLeuGlyAlaPheIleCysLysMetValProPheValGlnSerThrAla 128
Db 319 ACTGGTGGCCCTTCACAAATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCC 378
QY 129 ValValThrGluIleLeuThrMetThrCysIleAlaValGluArduIleGlnGlyLeuVal 148
Db 379 GTGTGAGTTGTTGTTTACATATGATGATGATGATGATGATGATGATGATGATGATGAT 438
QY 149 HisProPheCysMetLysTrpGlnIleThrAsnArgAlaPheThrMetLeuGlyVal 168
Db 439 CACCTTTTCGGGAGAG-----CTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
QY 169 ValTrpLeuValAlaValIleValGlySerProMetTrpHisValGlnGlnLeuGluIle 188
Db 493 ATCTGGCCCTGGCCCTGCTCATCAIGTCTCCCTCC-----GCCCTC 534
QY 189 LysTrpAspPheLeuTyrGluLysGluIleIle----- 199
Db 535 AGCTGACGCTCAACGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594
QY 200 -----CysCysLeuGluGluIleThrSerProValHisGlnLysIleIleThr 216
Db 595 CCTATCTATCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654

```

```

Db: 1048 GCCAACACCATCATCATAGCTACAGAGAAATTGCGGAGAGATTCCAGAGTACC 1107
US-09-255-368-7
: Sequence 7, Application US/09255368
: Patent No. 6252246
: GENERAL INFORMATION:
: APPLICANT: Gerald, Christophe P.G.
: APPLICANT: Jones, Kenneth A.
: APPLICANT: Bonini, James A.
: APPLICANT: Borowsky, Beth
: TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors
: TITLE OF INVENTION: and Uses Thereof
: FILE REFERENCE: 1795/57155-A
: CURRENT FILING DATE: 1999-02-22
: EARLIER APPLICATION NUMBER: 09/161,113
: EARLIER FILING DATE: 1998-09-25
: NUMBER OF SEQ ID NOS: 42
: SOFTWARE: Patent In Ver. 2.0 - beta
: SEQ ID NO 7
: LENGTH: 1293
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-255-368-7

Alignment Scores:
Pred. No.: 4,476-42 Length: 1293
Score: 432.50 Matches: 93
Percent Similarity: 50.32% Conservatives: 64
Best Local Similarity: 29.81% Mismatches: 126
Query Match: 19.35% Indels: 29
DB: 4 Caps: 5

US-10-070-241b-1 (1-431) x US-09-255-368-7 (1-1293)

QY 49 ValLeuThrGlyValLeuLeuPheAlaLeuPheGlyAsnAlaLeuValPheTyr 68
Db 139 ATGTGGCCATATGCCCTCATCTTCCTGCTCATGCTGGCCACACCTGCTGTTTC 198
QY 69 ValValThrArgSerLysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAla 88
Db 199 ATCGTGCTCAAGAACGGGACATCATATATATATATATATATATATATATATATAT 258
QY 89 LeuSerAspLeuLeuLeuThrPhePheCysIleProValThrMetLeuGlnAsnIleSer 108
Db 259 GTCAATGACCTGTGTGTGATATCTTGTGATATATATATATATATATATATATATAT 318
QY 109 AspAsnTrpLeuGlyAlaPheIleCysLysMetValProPheValGlnSerThrAla 128
Db 319 ACTGGTGGCCCTTCACAAATGCCATGCCATGCCATGCCATGCCATGCCATGCCATGCC 378
QY 129 ValValThrGluIleLeuThrMetThrCysIleAlaValGluArduIleGlnGlyLeuVal 148
Db 379 GTGTGAGTTGTTGTTTACATATGATGATGATGATGATGATGATGATGATGATGATGAT 438
QY 149 HisProPheCysMetLysTrpGlnIleThrAsnArgAlaPheThrMetLeuGlyVal 168
Db 439 CACCTTTTCGGGAGAG-----CTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
QY 169 ValTrpLeuValAlaValIleValGlySerProMetTrpHisValGlnGlnLeuGluIle 188
Db 493 ATCTGGCCCTGGCCCTGCTCATCAIGTCTCCCTCC-----GCCCTC 534
QY 189 LysTrpAspPheLeuTyrGluLysGluIleIle----- 199
Db 535 AGCTGACGCTCAACGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594
QY 200 -----CysCysLeuGluGluIleThrSerProValHisGlnLysIleIleThr 216
Db 595 CCTATCTATCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654

```











GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 06:53:16 ; Search time 6185 Seconds  
(without alignments)  
11363.505 Million cell updates/sec

Title: US-10-070-241B-2

Perfect score: 2415

Sequence: 1 gccagagccgcaggaccct. ....aaaaaaaaaaaaaaagg 2415

Scoring table: IDENTITY\_NUC  
Gapop 10 0 , Gapext 1 0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hq.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pi.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_of.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pi.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_hq\_hum.\*  
31: em\_hq\_inv.\*  
32: em\_hq\_other.\*  
33: em\_hq\_mus.\*  
34: em\_hq\_pln.\*  
35: em\_hq\_rod.\*  
36: em\_hq\_mam.\*  
37: em\_hq\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description         |
|------------|--------|---------------|--------|-------|---------------------|
| 1          | 1668.4 | 69.1          | 1710   | 6     | AX360197 Sequence   |
| 2          | 1296   | 53.7          | 1296   | 6     | AX360195 Sequence   |
| 3          | 1293   | 53.5          | 1293   | 6     | BD012995 Novel G p  |
| 4          | 1293   | 53.5          | 1293   | 23    | BD010028 Novel G p  |
| 5          | 1122.8 | 46.5          | 75164  | 9     | AC093867 Homo sapi  |
| 6          | 1121.2 | 46.4          | 155347 | 2     | AC005961 Homo sapi  |
| 7          | 1078.4 | 44.7          | 1368   | 9     | AF411117 Homo sapi  |
| 8          | 693.4  | 28.7          | 170227 | 9     | AC094816 Homo sapi  |
| 9          | 309    | 12.8          | 432    | 5     | AX277359 Sequence   |
| 10         | 309    | 12.8          | 432    | 6     | AX277395 Sequence   |
| 11         | 309    | 12.8          | 432    | 6     | AX354959 Sequence   |
| 12         | 278.8  | 11.5          | 24761  | 9     | AC005894 Homo sapi  |
| 13         | 278.2  | 11.5          | 256164 | 9     | HS4400877 Homo sapi |
| 14         | 273    | 11.3          | 170561 | 9     | AP005482 Homo sapi  |
| 15         | 273    | 11.3          | 193858 | 2     | AC007734 Homo sapi  |
| 16         | 273    | 11.3          | 199653 | 9     | AP001077 Homo sapi  |
| 17         | 270    | 11.2          | 168203 | 2     | AP001902 Homo sapi  |
| 18         | 269.8  | 11.2          | 6712   | 9     | ALR32112 Homo sapi  |
| 19         | 269.8  | 11.2          | 135038 | 9     | HSMR03419 Homo sapi |
| 20         | 268.2  | 11.1          | 171582 | 9     | HUMYX0703 Homo sapi |
| 21         | 264.4  | 10.9          | 126117 | 9     | AC021558 Homo sapi  |
| 22         | 263.6  | 10.9          | 15871  | 9     | AC004000 Human PAC  |
| 23         | 263.6  | 10.9          | 173803 | 9     | CNS01DV2 Human chr  |
| 24         | 263    | 10.9          | 180346 | 2     | AC115437 Human chr  |
| 25         | 261.4  | 10.8          | 183849 | 2     | AC115437 Rattus no  |
| 26         | 261.4  | 10.8          | 167556 | 2     | AC011219 Homo sapi  |
| 27         | 260.2  | 10.8          | 186624 | 2     | AC130878 Rattus no  |
| 28         | 259.2  | 10.7          | 173868 | 9     | CNS01DTN Human chr  |
| 29         | 258.6  | 10.7          | 161504 | 9     | AP001525 Homo sapi  |
| 30         | 258.6  | 10.7          | 210731 | 2     | AC006517 Homo sapi  |
| 31         | 257.8  | 10.7          | 215795 | 2     | AC124891 Homo sapi  |
| 32         | 257.4  | 10.7          | 196703 | 2     | AC127470 Pan trogl  |
| 33         | 257    | 10.6          | 47009  | 9     | AC096851 Pan trogl  |
| 34         | 256.4  | 10.6          | 44348  | 9     | ALB63061 Human DNA  |
| 35         | 256.2  | 10.6          | 174869 | 2     | AC006046 Homo sapi  |
| 36         | 256.2  | 10.6          | 198636 | 2     | AP001840 Homo sapi  |
| 37         | 256    | 10.6          | 173916 | 2     | AC019189 Homo sapi  |
| 38         | 255.8  | 10.6          | 94793  | 9     | AC068622 Homo sapi  |
| 39         | 255.6  | 10.6          | 59624  | 2     | AC006320 Homo sapi  |
| 40         | 255.6  | 10.6          | 100147 | 9     | AC091023 Homo sapi  |
| 41         | 255.6  | 10.6          | 132202 | 9     | AC011478 Homo sapi  |
| 42         | 255.6  | 10.6          | 133303 | 2     | AC025218 Homo sapi  |
| 43         | 255.6  | 10.6          | 199999 | 9     | AC05218 Homo sapi   |
| 44         | 255.4  | 10.6          | 138391 | 2     | AC004187 Homo sapi  |
| 45         | 255    | 10.6          | 137678 | 9     | AL357733 Homo sapi  |
|            |        |               |        |       | AL034351 Human DNA  |

# ALIGNMENTS

|            |  |                                   |     |        |                 |
|------------|--|-----------------------------------|-----|--------|-----------------|
| RESULT 1   | AX360197   | 1710 bp                           | DNA | linear | PAI 15-FEB-2002 |
| LOCUS      | AX360197   | Sequence 3 from Patent WO2004518. |     |        |                 |
| DEFINITION | AX360197   |                                   |     |        |                 |
| ACCESSION  | AX360197   |                                   |     |        |                 |
| VERSION    | AX360197.1   | GI:18675761                       |     |        |                 |
| KEYWORDS   | human.   |                                   |     |        |                 |
| SOURCE     | Human sapiens  |                                   |     |        |                 |
| ORGANISM   | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |                                   |     |        |                 |
| REFERENCE  | 1  |                                   |     |        |                 |
| AUTHORS    | Bloomquist, B.T. and Zhelnin, L.   |                                   |     |        |                 |
| TITLE      | Human neurotrophin Y-like 9 protein-coupled receptor   |                                   |     |        |                 |
| JOURNAL    | Patent: WO 00/04518-A 3 17-JAN-2002;   |                                   |     |        |                 |



ORIGIN

Query Match 53.7%; Score 1296; DB 6; Length 1296;  
Best Local Similarity 100.0%; Pred. No. 7.8e-255;  
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 ATGCAGGCGCTTAACATTACCCGAGCAGTTCTCTCGGCTGCTCGGGAGCACACCTG 413  
DB 1 ATGCAGGCGCTTAACATTACCCGAGCAGTTCTCTCGGCTGCTCGGGAGCACACCTG 60

QY 414 AGCGGGAGCAGTTATCGCTCTGTACCGGCTGCGACCGCTGCTACACCCAGAGCTG 473  
DB 61 AGCGGGAGCAGTTATCGCTCTGTACCGGCTGCGACCGCTGCTACACCCAGAGCTG 120

QY 474 CGGGAGCGCCCAAGCTGGCTGCTGCTACCGGCTGCTACCTTCGCTCGGCTGCTG 533  
DB 121 CGGGAGCGCCCAAGCTGGCTGCTGCTACCGGCTGCTACCTTCGCTCGGCTGCTG 180

QY 534 TTGSCAATGTTCTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 593  
DB 181 TTGSCAATGTTCTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 240

QY 594 AACATCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653  
DB 241 AACATCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

QY 654 GTCCATGCTCCAGAACATTTCCGACAACTGGCTGGGGTGGCTTTCATTTGCAAGTG 713  
DB 301 GTCCATGCTCCAGAACATTTCCGACAACTGGCTGGGGTGGCTTTCATTTGCAAGTG 360

QY 714 GTGCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 773  
DB 361 GTGCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 774 GTGGAAGCGCACAGGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 833  
DB 421 GTGGAAGCGCACAGGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 834 AGGGCTTTTACAACTGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 893  
DB 481 AGGGCTTTTACAACTGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 894 TGGCAGGCGCACAACTTGAGATCAATATGCTTCTTATATGAAAGAGCAATCTGCTG 953  
DB 541 TGGCAGGCGCACAACTTGAGATCAATATGCTTCTTATATGAAAGAGCAATCTGCTG 600

QY 954 TCGTTAGAAAGCTGGACAGCGCTGTCACCAAGAGATCTACACACCTTCTGCTGCTG 1013  
DB 601 TCGTTAGAAAGCTGGACAGCGCTGTCACCAAGAGATCTACACACCTTCTGCTGCTG 660

QY 1014 ATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073  
DB 661 ATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

QY 1074 CTTTGGATAAAGAAAGAGTTGGGATGGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1133  
DB 721 CTTTGGATAAAGAAAGAGTTGGGATGGTTGCTGCTGCTGCTGCTGCTGCTGCTG 780

QY 1134 ATGTCGCAAAATAGCCAGGAAGAAAGAAAGAGCTGCTGCTGCTGCTGCTGCTGCT 1193  
DB 781 ATGTCGCAAAATAGCCAGGAAGAAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 840

QY 1194 CTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1253  
DB 841 CTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

QY 1254 TTGGAAGAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313  
DB 901 TTGGAAGAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960

QY 1314 GGAATTTTCCAACTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1373  
DB 961 GGAATTTTCCAACTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

QY 1374 AAAAATGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1433  
DB 1021 AAAAATGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

QY 1434 AGGCATGGAATTTAGGAATTAATGATGTTGAAATGAAATGAAATGAAATGAAATG 1493  
DB 1081 AGGCATGGAATTTAGGAATTTAGGAATTTAGGAATTTAGGAATTTAGGAATTTAG 1140

QY 1494 AATCCATGAGGAGAAACCAAGGAGAGCAATTCAGTATGCGCAACATTTGAAGTCAAAT 1553  
DB 1141 AATCCATGAGGAGAAACCAAGGAGAGCAATTCAGTATGCGCAACATTTGAAGTCAAAT 1200

QY 1554 TGTCAACAGACAGAGGAG 1613  
DB 1201 TGTCAACAGACAGAGGAG 1260

QY 1614 CTGCTGAGAAATTTCTTTTATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1649  
DB 1261 CTGCTGAGAAATTTCTTTTATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1296

RESULT 3  
BD012995  
Locus  
DEFINITION Novel G protein coupled receptor protein and its DNA.  
ACCESSION BD012995  
VERSION BD012995.1 GI:22093184  
KEYWORDS WO 0116316-A/1.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 1293)  
AUTHORS Watanabe, F., Kikuchi, K. and Shintani, Y.  
TITLE Novel G protein coupled receptor protein and its DNA  
JOURNAL Patent: WO 0116316-A 1 08-MAR-2001;  
TAKEDA CHEMICAL INDUSTRIES LTD, TAKUYA WATANABE, KUNIKO KIKUCHI, YASUSHI SHINTANI  
COMMENT OS Homo sapiens (human)  
PN WO 0116316-A/1  
PD 08-MAR-2001  
PF 24-AUG-2000 WO 2000JP005684  
PR 27-AUG-1999 JP 99p 241530  
PI TAKUYA WATANABE, KUNIKO KIKUCHI, YASUSHI SHINTANI PC  
C12N15/12, C12N15/09, C07K14/705, C07K16/28, C12P21/02, A61K45/00, PC  
A61P43/00.  
CC G01N33/566, G01N33/15, G01N33/50

FEATURES  
source location/Qualifiers  
1..1293  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 342 a 305 c 301 g 345 t  
ORIGIN

Query Match 53.5%; Score 1293; DB 6; Length 1293;  
Best Local Similarity 100.0%; Pred. No. 3.2e-254;  
Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 ATGATGATGTTTAAATATTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 413  
DB 1 ATGATGATGTTTAAATATTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

QY 414 ACATGATGATGTTTAAATATTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473  
DB 61 ACATGATGATGTTTAAATATTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

QY 474 CCGGAG 533  
DB 121 CCGGAG 180

QY 534 TTGGAAG 593  
DB 181 TTGGAAG 240

QY 594 AACATCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653  
DB 241 AACATCTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

QY 654 GTCCATGCTCCAGAACATTTCCGACAACTGGCTGGGGTGGCTTTCATTTGCAAGTG 713  
DB 301 GTCCATGCTCCAGAACATTTCCGACAACTGGCTGGGGTGGCTTTCATTTGCAAGTG 360

QY 714 GTGCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 773  
DB 361 GTGCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 774 GTGGAAGCGCACAGGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 833  
DB 421 GTGGAAGCGCACAGGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 834 AGGGCTTTTACAACTGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 893  
DB 481 AGGGCTTTTACAACTGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 894 TGGCAGGCGCACAACTTGAGATCAATATGCTTCTTATATGAAAGAGCAATCTGCTG 953  
DB 541 TGGCAGGCGCACAACTTGAGATCAATATGCTTCTTATATGAAAGAGCAATCTGCTG 600

QY 954 TCGTTAGAAAGCTGGACAGCGCTGTCACCAAGAGATCTACACACCTTCTGCTGCTG 1013  
DB 601 TCGTTAGAAAGCTGGACAGCGCTGTCACCAAGAGATCTACACACCTTCTGCTGCTG 660

QY 1014 ATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073  
DB 661 ATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

QY 1074 CTTTGGATAAAGAAAGAGTTGGGATGGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1133  
DB 721 CTTTGGATAAAGAAAGAGTTGGGATGGTTGCTGCTGCTGCTGCTGCTGCTGCTG 780

QY 1134 ATGTCGCAAAATAGCCAGGAAGAAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 1193  
DB 781 ATGTCGCAAAATAGCCAGGAAGAAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 840

QY 1194 CTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1253  
DB 841 CTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

QY 1254 TTGGAAGAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313  
DB 901 TTGGAAGAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960

QY 1314 GGAATTTTCCAACTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1373  
DB 961 GGAATTTTCCAACTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020











|               |                          |                   |               |                   |                            |
|---------------|--------------------------|-------------------|---------------|-------------------|----------------------------|
| repeat_region | 23880..24560             | /rpt_family="L1"  | repeat_region | /rpt_family="Alu" | 92523..92927               |
| repeat_region | 26241..26519             | /rpt_family="Alu" | repeat_region | /rpt_family="Alu" | 92932..93218               |
| repeat_region | 26646..26920             | /rpt_family="Alu" | repeat_region | /rpt_family="Alu" | 98713..98971               |
| repeat_region | 27478..27760             | /rpt_family="Alu" | repeat_region | /rpt_family="Alu" | complement(101354..101559) |
| repeat_region | 28463..28745             | /rpt_family="Alu" | repeat_region | /rpt_family="Alu" | complement(101576..101683) |
| repeat_region | 28924..29159             | /rpt_family="Alu" | repeat_region | /rpt_family="Alu" | 102229..102552             |
| repeat_region | 29659..29774             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | 102558..102665             |
| repeat_region | 30064..30219             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(104177..104560) |
| repeat_region | 31258..31368             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(105291..110361) |
| repeat_region | complement(34992..40282) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(108878..109028) |
| repeat_region | complement(40832..40942) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(110483..110652) |
| repeat_region | 42058..42689             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | 112817..113106             |
| repeat_region | 42745..42794             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(113723..113889) |
| repeat_region | complement(46638..47151) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(116719..116785) |
| repeat_region | 47293..47351             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | 116775..116946             |
| repeat_region | complement(48072..49022) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(120455..120511) |
| repeat_region | complement(49129..49432) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | 123993..124318             |
| repeat_region | 50485..50683             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(124810..125066) |
| repeat_region | complement(51366..51451) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(129701..130088) |
| repeat_region | 57160..57424             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | 134127..135016             |
| repeat_region | 57619..57917             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | 137132..138120             |
| repeat_region | complement(57986..58102) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(137269..137558) |
| repeat_region | complement(59941..60025) | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | complement(139196..139483) |
| repeat_region | 60057..60352             | /rpt_family="L1"  | repeat_region | /rpt_family="L1"  | 140097..140362             |
| repeat_region | 63695..63963             | /rpt_family="L1"  |               |                   |                            |
| repeat_region | 67457..67895             | /rpt_family="L1"  |               |                   |                            |
| repeat_region | complement(69113..69482) | /rpt_family="L1"  |               |                   |                            |
| repeat_region | complement(69879..70176) | /rpt_family="L1"  |               |                   |                            |
| repeat_region | 70985..71148             | /rpt_family="L1"  |               |                   |                            |
| repeat_region | 71150..71472             | /rpt_family="L1"  |               |                   |                            |
| repeat_region | 71478..71963             | /rpt_family="L1"  |               |                   |                            |
| repeat_region | complement(72510..72712) | /rpt_family="L1"  |               |                   |                            |
| repeat_region | 76772..77064             | /rpt_family="L1"  |               |                   |                            |
| repeat_region | complement(81969..82102) | /rpt_family="L1"  |               |                   |                            |
| repeat_region | 82246..82490             | /rpt_family="L1"  |               |                   |                            |
| repeat_region | complement(86124..86418) | /rpt_family="L1"  |               |                   |                            |
| repeat_region | 87429..87692             | /rpt_family="L1"  |               |                   |                            |

Query Match 46.4%; Score 1121.2; DB 2; Length 155347;  
 Best Local Similarity 99.0%; Pred. No. 1.2e-218;

Matches 1126; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1249 GTAATTTTCAAAGGAATATGATGTCACAAATCAAGATGATTTTGTCTATCGTGGAAA 1308

DB 59887 GTAATTTTCAAAGGAATATGATGTCACAAATCAAGATGATTTTGTCTATCGTGGAAA 59946

QY 1309 TTATTGGATTTTCCAACTCCATCTGTAATCCAAATGTCATGCAATTTATGAATGAAACT 1368

DB 59947 TTATTGGATTTTCCAACTCCATCTGTAATCCAAATGTCATGCAATTTATGAATGAAACT 60006

QY 1369 TCAAAAATAATGTTTGTCTGAGTTTGTATTGATGATGATTAATAAATGAACTTCTCTCCAG 1428

DB 60007 TCAAAAATAATGTTTGTCTGAGTTTGTATTGATGATGATTAATAAATGAACTTCTCTCCAG 60066

QY 1429 CACAAAGGATGGAATATGAGGAATATGAGGAATATGAGGAATGAGGAATGAGGAATGAGGAAT 1488

DB 60067 CACAAAGGATGGAATATGAGGAATATGAGGAATATGAGGAATGAGGAATGAGGAATGAGGAAT 60126

QY 1489 GAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548

DB 60127 GAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60186

[illegible]











Viel, R., Vo, A., Willson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**REFERENCE 4** (bases 1 to 224761)  
**JOURNAL:** Submitted (13-AUG-2002) Whitehead Institute/MIT Center for Genome  
**TITLE:** Direct Submission  
**AUTHOR(S):** Zhang L.H. and Boyan M.

4. (bases 1) (O 224/761)

Birren, B., Nisbun, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boquslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArcangelo, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-pierre, N., Haqos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Melidrim, J., Menous, I., Mikova, T., Mlenda, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Reita, R., Rise, C., Rodov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
JOURNAL  
Zimmer, G., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 21, 2002 this sequence version replaced gi:22213414.

on Aug 21, 2002. This sequence version replaced g1:22213419.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

//www.genome.washington.edu/gen/submitMasker.html  
 Genomic Center  
 Center, Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIGR  
 Web site: <http://www.ssq.wi.mit.edu>  
 Contact: sequence.submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L6446  
 Center clone name: 318 C 2

```

FEATURES
  source
    location/Qualifiers
      1. .224761
        /orqanism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="RP11-41BC2"
        /clone_lib="RP11-41BC2"
        repeat_region="complement (992..1290)"
        Male_BAC=""

```

|               |                           |               |
|---------------|---------------------------|---------------|
| repeat_region | /rpt_family="L3"          | 11008, .11295 |
| repeat_region | /rpt_family="Aluv"        | 11296, .11411 |
| repeat_region | /rpt_family="L3"          | 11801, .12035 |
| repeat_region | /rpt_family="Alud"        | 12036, .12185 |
| repeat_region | complement(12401, .12685) |               |
| repeat_region | /rpt_family="Alusx"       | 13447, .13498 |
| repeat_region | /rpt_family="(TA)n"       | 14889, .14921 |
| repeat_region | complement(13506, .13787) |               |
| repeat_region | /rpt_family="Alusx"       | 14218, .14518 |
| repeat_region | /rpt_family="Alusx"       | 14537, .14875 |
| repeat_region | /rpt_family="HAL1"        | 14889, .14921 |
| repeat_region | /rpt_family="(TTCA)n"     | 15312, .15312 |
| repeat_region | /rpt_family="Alud"        | 15369, .16211 |
| repeat_region | /rpt_family="L2"          | 16212, .16484 |
| repeat_region | /rpt_family="MER38B"      | 16492, .16646 |
| repeat_region | /rpt_family="L2"          | 16661, .16800 |
| repeat_region | /rpt_family="L2"          | 16867, .17148 |
| repeat_region | /rpt_family="Alusq"       | 17277, .17477 |
| repeat_region | /rpt_family="MER112"      | 17522, .17788 |
| repeat_region | /rpt_family="Alusx"       | 17801, .18020 |
| repeat_region | /rpt_family="HAL1"        | 18048, .18456 |
| repeat_region | /rpt_family="LIMA7"       | 18459, .18495 |
| repeat_region | /rpt_family="AT-rich"     | 18631, .18868 |
| repeat_region | /rpt_family="HAL1"        | 18902, .19261 |
| repeat_region | /rpt_family="HAL1"        | 19324, .19635 |
| repeat_region | /rpt_family="Alusq"       | 19636, .19890 |
| repeat_region | /rpt_family="Alusq"       | 21016, .21071 |
| repeat_region | /rpt_family="L2"          | 21482, .21704 |
| repeat_region | /rpt_family="MER33"       | 22258, .22587 |
| repeat_region | /rpt_family="Alusx"       | 22636, .22712 |
| repeat_region | /rpt_family="MIR3"        | 24907, .25030 |
| repeat_region | /rpt_family="L2"          | 27043, .27345 |
| repeat_region | /rpt_family="Alusx"       | 27586, .27894 |
| repeat_region | /rpt_family="Alusx"       | 28617, .28864 |
| repeat_region | /rpt_family="Alud"        | 29157, .29455 |
| repeat_region | /rpt_family="Alusp"       | 29555, .29846 |
| repeat_region | /rpt_family="Alusq"       | 30152, .30152 |
| repeat_region | /rpt_family="Alusp"       | 30161, .30409 |

|                           |  |        |
|---------------------------|--|--------|
| Query Match               | 11.5%; Score 278.8; DB 9; Length 224761;   | source |
| Best Local Similarity     | 71.7%; Pred. No. 2.6e-46;  |        |
| Matches 393; Conservative | 0; Mismatches 149; Indels 6; Gaps 2;   |        |
| QY 1678                   | CTTCAGATTTGACCCAGAGAAATATTTTGGACAAAGGTCAAAATCTCTTTTAATT 1737   |        |
| DB 110394                 | CTGAGATAATTCATTAATCTAAGACTGATACCATTAGAACAAAGATTTGGTAACCTAGTG 110453  |        |
| QY 1738                   | CTTAAGATGATGACAGAGAAACAAATCATCTTTCCATTAAAAATGACACGAGGCTA 1797  | gene   |
| DB 110454                 | TTTCAAGAGTAAATTTTTTAGACTTTAATTTCTCCATAAAAAATATCTTACAGGCTG 110513   | mRNA   |
| QY 1798                   | GTCCAACTGGAGTGATGTTTACAAACCAATNGATCACAATCATTTAACANATTTCTGTGPT 1857   |        |
| DB 110514                 | GTTCAAGTGTAGTATGTTTACAACTAATTTGATGACAAACAGTTAAGATTTCTTTGTTT 110573   |        |
| QY 1858                   | CTTCTCATTTCCCACTGCTTTCATCTGACGTAGCTTTAAAGAGCAACATGGAAGCCAGGC 1917  |        |
| DB 110574                 | CTTCTGCTCCCACTGCTTTCATTTCACTAGTCTTTTAAAAAATAAAAA---AAGGCGGGC 110630  | exon   |
| QY 1918                   | ACGTGGCTCATGCTGTAAATCCCACTACCTTTGAGAGAGCTAGAGGAGGATCAAGAG 1977   | CDS    |
| DB 110631                 | ACGATGCTCAGGCTGTATCCCACTTTTGGAGAGGCAAGGCGGTTGGATCACAAGG 110690   |        |
| QY 1978                   | TCAGGAGATCAAAACCACTCTGCTAACAGGTGAACCCCACTCTGCTGCTAAAAATAC-- 2035   |        |
| DB 110691                 | TCAGGAGATTGAGACCACTCTGCTAACATGTTGAACCTGTTTCTACTAAAAATACAA 110750   |        |
| QY 2036                   | AAAAATAGCGGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2094   |        |
| DB 110751                 | AAAAATAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 110810   |        |
| QY 2095                   | GGGAGAAATGTTGAACCGGAGGCGGAGCTTGCAGTGATCGGAGATCATGCCACTGC 2154  |        |
| DB 110811                 | CAGGAGAAATGCAATGAACCGGAGGCTGAGCTTGCAGTGAGCGAGATCACCACCTGC 110870   |        |
| QY 2155                   | ACTCCAGCTGGGGAAGACGACGACTCCCGCTCTCAAAAAAATTTTTTTTCAAAAAAT 2214   |        |
| DB 110871                 | ACTCCAGCTGGGGAAGACGAGAGCTCGCTCTCAAAAAAATTTTTTTTCAAAAAAAT 110930  |        |
| QY 2215                   | CGTAAACC 2222  |        |
| DB 110931                 | CTTCAAGC 110938  |        |
| RESULT 13                 |  |        |
| HSA400877                 |  |        |
| LOCUS                     |  |        |
| DEFINITION                | Homo sapiens ASCL3 gene, CEGP1 gene, Cllorlf14 gene, Cllorlf15 gene, Cllorlf16 gene and Cllorlf17 gene.  |        |
| ACCESSION                 | AJ400877   |        |
| VERSION                   | AJ400877.1 GI:8052236  |        |
| KEYWORDS                  | ASCL3 gene; Cllorlf14 gene; Cllorlf15 gene; Cllorlf16 gene; Cllorlf17 gene; CEGP1 gene.  |        |
| SOURCE                    | human.   |        |
| ORGANISM                  | Homo sapiens   |        |
| REFERENCE                 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.   |        |
| AUTHORS                   | 1 Amid, C., Bahr, A., Mujica, A., Sampson, N., Biker, S.E., Winterpacht, A., Zabel, R., Hankeln, T. and Schmidt, R.P.  |        |
| TITLE                     | Comparative genomic sequencing reveals a strikingly similar architecture of a conserved syntenic region on human chromosome 1p15.3 (including gene ST5) and mouse chromosome 7 |        |
| JOURNAL                   | Cytogenet. Cell Genet. 93 (3-4), 284-290 (2001)  |        |
| MEDLINE                   | 21418998   |        |
| PUBMED                    | 11528127   |        |
| REFERENCE                 | 2 (bases 1 to 256164)  |        |
| AUTHORS                   | Hankeln, T.  |        |
| TITLE                     | Direct Submission  |        |
| JOURNAL                   | Submitted (27-APR-2000) Hankeln T., Inst. Molekulargenetik, Univ. Mainz, Becherweg 32, D-55099 Mainz, GERMANY  |        |
| FEATURES                  | Location/Qualifiers  |        |

```

1. .256164
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p15.3"
/clone="LNL704J141117Q3"
/clone="PAC275G1"
/clone_lib="LNL704"
18454..89659
/gene="CEGP1"
join(18454..18666,20233..20355,30545..30670,35439..35573,
40559..40684,43241..43357,44074..44163,49530..49646,
50629..50751,54145..54261,56293..56415,56837..56953,
59338..59442,61950..62105,62486..62692,76251..76421,
79122..79289,80002..80163,82486..82683,84193..84345,
88092..88171,88850..89659)
/gene="CEGP1"
/evidence="experimental"
18454..18666
/gene="CEGP1"
/number=1
join(18534..18666,20233..20355,30545..30670,35439..35573,
40559..40684,43241..43357,44074..44163,49530..49646,
50629..50751,54145..54261,56293..56415,56837..56953,
59338..59442,61950..62105,62486..62692,76251..76421,
79122..79289,80002..80163,82486..82683,84193..84345,
88092..88171,88850..89002)
/gene="CEGP1"
/codon_start=1
/product="CEGP1 protein"
/protein_id="CAB92285.1"
/db_xref="GI:8052237"
/db_xref="SPTREMBL:Q9NGA6"
/translation="MGVAGNRNPQAAWVLLLLLLPPLLLAGAVPPGRGAAQPOE
DVEPCAGLLDHAVALVNTPTSYKSPKPYQKPGQFRIIDCGNELNGQVHC
LNIPGNYRCTCFGLAHADGHNCLDVECLNNGCQCTCVNMGSTECCKGQFEL
SDNQHICIRHSEGLSKMKNHCCSHICEAPRGVACECPGFELAKNQCILUTN
HGNGGQCHSDDTADGPECCHPOYKMTDGRSCLEDEDVLEVTSTTTSVVDGDKR
VKRILMETCAVANGGCDRTKDTSTGTHGCSPPVGFILQLDKTKCIDEGCTRNNGC
DHFCKTGVSGDCGCKKFKLLTDEKSCQDVECSLDFTCDHSCINHPGTFACACNRG
YILYKTHGDI NECSINNGGQUCVNTVGSYECQCHPGYKLIHWNKDAVEKGLLP
TSVSPVSLHCGSGGGDCGFLRCHSGIHLSSDVTITRTSVFKINECKCSIKNAELP
PEGRPALPEKSSVAFSPRYNLTCSCKQVPGAPRSTPKEMFTIPEFELETNOK
EVTASCDLSICVKRTEKRLKALITLRAVHRQPHLOISGMNIDVAKKPKPTSERQA
ESGCVQHAENAVSRRASTYTKAFERILDPNITFQNEEQTDECPGPGNSGA
LKTPEANNMSECGGLQPGYSALAFAPQLALSTFQPEASPTSCFPGGLATKHQ
GATSFDAETRVQSPGFHYNTTHRCIRCPVSTYQPEGKNVVSQPGNTTDFDGS
TNLTQCKRCCGGELODFTCYIESPNYPONYPANTECTWTINPPKRLILVPEIFEL
PIEDDGDYLVMRKTSNNSVTITCTQTERPIAFTSRKKLITQFKNPNSGARGF
QVPYVYDEYQELIHDIVRDQKIVASENHQELIKKKLILKALPDLVLAHPONYPKYTA
QESPMPFPSPFLPLPSKVSFLRPYK"
18667..20232
/gene="CEGP1"
/number=1
20233..20355
/gene="CEGP1"
/number=2
20356..30544
/gene="CEGP1"
/number=2
30545..30670
/gene="CEGP1"
/number=3
30671..35438
/gene="CEGP1"
/number=3
35439..35573
/gene="CEGP1"
/number=4
35574..40558
/gene="CEGP1"
/number=4
40559..40684

```

```

/ gene- "CEGP1"
/ number-5
40685..43240
/ gene- "CEGP1"
/ number-5
43241..43357
/ gene- "CEGP1"
/ number-6
43358..44073
/ gene- "CEGP1"
/ number-6
44074..44163
/ gene- "CEGP1"
/ number-7
44164..44929
/ gene- "CEGP1"
/ number-7
44930..44964
/ gene- "CEGP1"
/ number-8
44967..45062
/ gene- "CEGP1"
/ number-8
45069..45075
/ gene- "CEGP1"
/ number-9
45075..45144
/ gene- "CEGP1"
/ number-9
45145..45261
/ gene- "CEGP1"
/ number-10
45262..45292
/ gene- "CEGP1"
/ number-10
45293..45615
/ gene- "CEGP1"
/ number-11
45616..45683
/ gene- "CEGP1"
/ number-11
45687..45693
/ gene- "CEGP1"
/ number-12
45694..45937
/ gene- "CEGP1"
/ number-12
45938..45943
/ gene- "CEGP1"
/ number-13
45944..46199
/ gene- "CEGP1"
/ number-13
46195..46210
/ gene- "CEGP1"
/ number-14
46210..46248
/ gene- "CEGP1"
/ number-14
46248..46269
/ gene- "CEGP1"
/ number-15
46269..46270
/ gene- "CEGP1"
/ number-15
46271..46421
/ gene- "CEGP1"
/ number-16
46422..46421
/ gene- "CEGP1"
/ number-16
46422..46429
/ gene- "CEGP1"

```

```

intron
79290..80001
/ gene- "CEGP1"
/ number-17
80002..80163
/ gene- "CEGP1"
/ number-18
80164..82485
/ gene- "CEGP1"
/ number-18
82486..82683
/ gene- "CEGP1"
/ number-19
82684..84192
/ gene- "CEGP1"
/ number-19
84193..84345
/ gene- "CEGP1"
/ number-20
84346..88091
/ gene- "CEGP1"
/ number-20
88092..88171
/ gene- "CEGP1"
/ number-21
88172..88849
/ gene- "CEGP1"
/ number-21

```

```

Query Match 11.5%; Score 278.2; DB 9; Length 256164;
Best Local Similarity 73.1%; Pred. No. 3.5e-46;
Matches 384; Conservative 0; Mismatches 135; Indels 6; Gaps 2;

QY 1678 CTTCAATGTAAACCAAGAGAGAAATATTTGAGCAAGGTCACATACCTTTTAATT 1737
DB 232259 CTGGAATATATTCATTAAATCTAGACATGATATTAACAAAGATGTAACTAGTG 232318
QY 1738 CTTACATGATGACCAAGAGAGAAATATTCATGTTTCCATTAATAAATGACAGAGGTA 1797
DB 232319 TTCAATGAGTAAATTTTACACTTAAATTTCTCCAAATAAAAAATATCTTCAGGCTG 232378
QY 1798 GTCCAAAGTCAGTATGTTTACCAAGCAATGATGACAAATCATATTAACAAATTTCTGTTT 1857
DB 232379 GTTCAAGTCAGTATGTTTACCAAGCAATGATGACAAATCATATTAACAAATTTCTGTTT 232438
QY 1858 CCTTCTCATTCCTCCAGTCTTACACTTACCTTAAAGCAAGCAACATGAGGCGAGGC 1917
DB 232439 CTTCCTGCTGCAATGCTTCAATTAATTTTCAATTTATTTTAAATAAAAAAATAAAGGCGGCG 232495
QY 1918 AAGTGGTCTGATGCTTCAATTAATTTTCAATTTATTTTCAATTTATTTTCAATTTATTTT 1977
DB 232496 AGCATGGCTGAGGCTGATTCCTCCAGCCTTGGAGGCGGAGGCGGCTGATCACAAGG 232555
QY 1978 TCAGAGATGCAAAATATCTGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 2035
DB 232556 TCAGAGATGCAAAATATCTGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 232615
QY 2036 AAAAAATTAGCGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2094
DB 232616 AAAAAATTAGCGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 232675
QY 2095 GAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2154
DB 232676 GAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 232735
QY 2155 AATTCAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2199
DB 232736 AATTCAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 232780

```

RESULT 14  
AF005482  
Locus

AP005482 170561 bp DNA linear PRI 10-AUG-2002

| LOCUS<br>DEFINITION | ACCESSION<br>VERSION<br>KEYWORDS<br>SOURCE<br>ORGANISM | REFERENCE<br>AUTHORS<br>TITLE<br>JOURNAL<br>REFERENCE<br>AUTHORS | TITLE<br>JOURNAL<br>COMMENT |
|---------------------|--|--|-----------------------------|
|                     |  |  |                             |

AC007734 193858 bp dNA linear HTC 08-DPC-2007  
Homo sapiens chromosome 18 clone RP11-4401 map 18, WORKING DRAFT  
SEQUENCE, 23 unordered pieces.  
AC007734

AC007734 4 GI:11225456  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 193858)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 18, clone RP11-4401  
2 (bases 1 to 193858)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,  
Castle,A., Cerny,J., Collangelo,M., Collins,S., Collymore,A.,  
Cooke,P., DeArelano,K., Delapierre,B., Devon,K., Dewar,K.,  
Donohue,D., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,  
Fuks,B., Gage-B., Galadon,J., Gardyna,S., Gilbert,D., Grant,G.,  
Haags,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,  
Karakalas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,  
Matukas,N., McManus,P., McGuirk,A., McKernan,K., McLaughlin,J.,  
Melchior,J., Molla,M., Morris,W., Morrow,J., Mychalockyj,J.,  
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,  
Peterson,K., Pollara,J., Riley,P., Roberts,D., Roy,A., Severy,P.,  
Stasenko-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,  
Tsaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,  
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.

Direct Submission  
Submitted (05-JUN-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
on Nov 18, 2000 This sequence version replaced g1:8112158.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/HM/RepeatMasker.html>

----- Genome Center -----  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information -----  
Center project name: I683  
Center clone name: 44\_0.1  
----- Summary Statistics -----  
Sequencing vector: M13, M7815, 90% of reads  
Sequencing vector: Plasmid: L08752; 4% of reads  
Sequencing vector: Plasmid: n/a; 6% of reads  
Chemistry: Dye-primer-amersham; 80% of reads  
Chemistry: Dye-terminator Big Eye; 20% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 181500 bases at least Q40  
Consensus quality: 186874 bases at least Q30  
Consensus quality: 189522 bases at least Q20  
Insert size: 190000; agarose-gel  
Insert size: 191658; sum.  
NOTE: This is a 'working draft' sequence. It currently  
consists of 23 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 4108: contig of 4108 bp in length  
\* 4109 4208: gap of 100 bp  
\* 4209 4602: contig of 394 bp in length  
\* 4603 4702: gap of 100 bp  
\* 4703 5492: contig of 790 bp in length  
\* 5493 5592: gap of 100 bp  
\* 5593 6962: contig of 1370 bp in length  
\* 6963 7062: gap of 100 bp

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 23 contigs, the true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence,  
\* as soon as it is available and the accession number will  
\* be preserved.

```

* 7064 7987: contig of 925 bp in length
* 7988 8087: gap of 100 bp
* 8088 8966: contig of 879 bp in length
* 8967 9066: gap of 100 bp
* 9067 11029: contig of 1963 bp in length
* 11030 11129: gap of 100 bp
* 11130 12990: contig of 1861 bp in length
* 12991 13090: gap of 100 bp
* 13091 14979: contig of 1889 bp in length
* 14980 15079: gap of 100 bp
* 15080 16198: contig of 1119 bp in length
* 16199 16298: gap of 100 bp
* 16299 18878: contig of 2580 bp in length
* 18879 18978: gap of 100 bp
* 18979 22036: contig of 3058 bp in length
* 22037 22136: gap of 100 bp
* 22137 25223: contig of 3087 bp in length
* 25224 25323: gap of 100 bp
* 25324 29254: contig of 3931 bp in length
* 29255 29354: gap of 100 bp
* 29355 35127: contig of 5973 bp in length
* 35128 35427: gap of 100 bp
* 35428 41718: contig of 6291 bp in length
* 41719 41818: gap of 100 bp
* 41819 47796: contig of 5978 bp in length
* 47797 47896: gap of 100 bp
* 47897 58805: contig of 10909 bp in length
* 58806 58905: gap of 100 bp
* 58906 104252: contig of 45347 bp in length
* 104253 104352: gap of 100 bp
* 104353 121266: contig of 16914 bp in length
* 121267 121366: gap of 100 bp
* 121367 141580: contig of 20214 bp in length
* 141581 141680: gap of 100 bp
* 141681 177437: contig of 35757 bp in length
* 177438 177537: gap of 100 bp
* 177538 194858: contig of 16321 bp in length.
FEATURES
    source
        1..194858
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="18"
            /map="18"
            /clone="RP11 4401"
            /clone_lib="RP11 Human Male BAC"
        1..4108
            /note="assembly_fragment"
            /clone_end:3p6
            /vector_side:left
            4209..4602
                /note="assembly_fragment"
                4703..5492
                    /note="assembly_fragment"
                    5593..6962
                        /note="assembly_fragment"
                        7064..7987
                            /note="assembly_fragment"
                            8088..8966
                                /note="assembly_fragment"
                                9067..11029
                                    /note="assembly_fragment"
                                    11130..12990
                                        /note="assembly_fragment"
                                        13091..14979
                                            /note="assembly_fragment"
                                            15080..16198
                                                /note="assembly_fragment"
                                                16299..18878
                                                    /note="assembly_fragment"
                                                    18979..22036
                                                        /note="assembly_fragment"
                                                        22137..25223
                                                            /note="assembly_fragment"

```

```

misc_feature 25324..29254
            /note="assembly_fragment"
misc_feature 29355..35327
            /note="assembly_fragment"
misc_feature 35428..41718
            /note="assembly_fragment"
misc_feature 41819..47796
            /note="assembly_fragment"
misc_feature 47897..58805
            /note="assembly_fragment"
misc_feature 58906..104252
            /note="assembly_fragment"
misc_feature 104353..121266
            /note="assembly_fragment"
misc_feature 121367..141580
            /note="assembly_fragment"
misc_feature 141681..177437
            /note="assembly_fragment"
misc_feature 177538..193858
            /note="assembly_fragment"
            clone_end:T7
            vector_side:right
BASE COUNT 54379 a 40835 c 40675 g 55766 t 2203 others
ORIGIN
Query Match 11.3% Score 273; DB 2; Length 193858;
Best Local Similarity 78.7%; Pred. No. 3.9c 45;
Matches 355; Conservative 0; Mismatches 82; Indels 14; Gaps 2;
QY 1776 ATTAATAAATGACACGAGGCTAGTCCATNTGAGTAGTATGTTTACACCAATNATACAA 1835
DB 109122 ATACATAAAGGACAGGTTGTGAGATGAGTGTGATTTTACAAATTAATTATGACAA 109181
QY 1836 TCATTAAACANATTTCTGTTCTCTCTCATTCCTCCACTGCTTACCTGACTAG----- 1888
DB 109182 TCAGTTACAGATTTTGTGTTCTTTTCAATTCATTAATTAATTAATTAATTAATTA 109241
QY 1889 -----CCTTAATAAAGGACACATGAGAGGACAGGAGGATATAGATATATATATATAT 1942
DB 109242 CATTTTATTAATAAAGACAGAGCTATAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 109301
QY 1943 GATTTTGTGAGGCTAGACAGGAGGATATAGATATATATATATATATATATATATATAT 2092
DB 109302 GATTTTGTGAGGCGCGAGGAGGATATATAGATATATATATATATATATATATATAT 109361
QY 2003 AACAGGTGAACCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2061
DB 109362 AACAGGTGAACCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 109421
QY 2062 CACTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2121
DB 109422 TGGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 109481
QY 2122 GATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2181
DB 109482 ATAGCTTGTATATATATATATATATATATATATATATATATATATATATATATAT 109541
QY 2182 TCCCGCTCTCAAAAAAATTTTTCGAAAAA 2212
DB 109542 TCCGCTCTCGGAAAAAATTTTTCGAAAAA 109572

```

Search completed: June 17, 2003, 10:08:15

Job time : 6213 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 06:16:42 ; Search time 1348 seconds

(without alignments)  
4034.552 Million cell updates/ser

Title: US-10-070-241b-2

Perfect score: 2415

Sequence: 1 gccagagggcgcaggagcct... aaaaaaaaaaaaaaaaaagg 2415

Scoring table: IDENTITY\_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N\_Geneseq\_101002.4

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARY

| Result No. | Score  | Query Match | Length | ID | Description        |
|------------|--------|-------------|--------|----|--------------------|
| 1          | 2412   | 99.9        | 2415   | 22 | Human G protein co |
| 2          | 2389.6 | 98.9        | 2411   | 21 | cdna encoding huma |
| 3          | 1668.4 | 69.1        | 1710   | 24 | Human neurotrophin |
| 4          | 1296   | 53.7        | 1296   | 22 | Human GTP binding  |
| 5          | 1296   | 53.7        | 1296   | 22 | Human G protein co |
| 6          | 1296   | 53.7        | 1296   | 24 | Human neurotrophin |
| 7          | 1296   | 53.7        | 1296   | 24 | Human neurotrophin |
| 8          | 1294.4 | 53.6        | 1296   | 21 | Human orphan G pro |
| 9          | 1294.4 | 53.6        | 1296   | 21 | Human G protein co |

|    |        |      |        |    |          |                    |
|----|--------|------|--------|----|----------|--------------------|
| 10 | 1293   | 53.5 | 1293   | 22 | AAF81818 | Human G protein co |
| 11 | 1289.6 | 53.4 | 1296   | 21 | AAA46115 | Human G protein co |
| 12 | 994.6  | 41.2 | 1737   | 24 | ABK86292 | Mouse TGR346b cDNA |
| 13 | 932    | 38.6 | 1772   | 24 | ABK86293 | Mouse TGR346b cDNA |
| 14 | 737    | 30.5 | 741    | 24 | AAS98050 | Human DNA for pote |
| 15 | 500    | 20.7 | 511    | 21 | AAD01143 | Human orphan G pro |
| 16 | 500    | 20.7 | 511    | 21 | AAA46044 | Human G protein co |
| 17 | 309    | 12.8 | 432    | 22 | AA170202 | Human GPCR-like pr |
| 18 | 309    | 12.8 | 432    | 24 | ABL42350 | Human galanin rece |
| 19 | 309    | 12.8 | 432    | 24 | AAS98092 | Human DNA for pote |
| 20 | 251.4  | 10.4 | 114793 | 22 | AAD08215 | Human genome from  |
| 21 | 251.2  | 10.4 | 506    | 22 | ABA15628 | Human nervous syst |
| 22 | 251    | 10.4 | 3178   | 22 | AAK80500 | Human immune/haema |
| 23 | 249.8  | 10.3 | 501    | 22 | ABA15627 | Human nervous syst |
| 24 | 249.8  | 10.3 | 31871  | 23 | ABK42516 | Genomic sequence # |
| 25 | 249.8  | 10.3 | 40742  | 22 | ABK68089 | Human immune/haema |
| 26 | 249.8  | 10.3 | 40742  | 22 | AAK79886 | Human immune/haema |
| 27 | 249.8  | 10.3 | 220895 | 24 | ABK84798 | Human cDNA differe |
| 28 | 249.4  | 10.3 | 3178   | 22 | AAK80499 | Human immune/haema |
| 29 | 249.2  | 10.3 | 267156 | 24 | ABK68560 | Kidney cancer rela |
| 30 | 249    | 10.3 | 333    | 22 | AAK66611 | Human immune/haema |
| 31 | 249    | 10.3 | 333    | 22 | AAK66612 | Human immune/haema |
| 32 | 249    | 10.3 | 333    | 22 | AAK66613 | Human immune/haema |
| 33 | 248    | 10.3 | 331    | 22 | AAK78782 | Human immune/haema |
| 34 | 248    | 10.3 | 4124   | 22 | AAS01154 | 5'-untranslated re |
| 35 | 248    | 10.3 | 32145  | 22 | AAK68491 | Human immune/haema |
| 36 | 248    | 10.3 | 32145  | 22 | AAK68575 | Human immune/haema |
| 37 | 248    | 10.3 | 144460 | 21 | AA293815 | Olfactory receptor |
| 38 | 247.8  | 10.3 | 31474  | 22 | AAK05461 | Human reproductive |
| 39 | 247.8  | 10.3 | 31474  | 23 | ABL98314 | Human testicular a |
| 40 | 247.4  | 10.2 | 11360  | 22 | AAK03820 | Human reproductive |
| 41 | 247.4  | 10.2 | 11360  | 22 | AAK04727 | Human reproductive |
| 42 | 246.8  | 10.2 | 13868  | 22 | AAK77131 | Human immune/haema |
| 43 | 246.4  | 10.2 | 314    | 22 | AAK76575 | Human immune/haema |
| 44 | 246.4  | 10.2 | 327    | 22 | AAK74122 | Human immune/haema |
| 45 | 246.4  | 10.2 | 24964  | 24 | ABK86595 | Human SA (rat type |

## ALIGNMENTS

RESULT 1  
AAF81818  
ID AAF81819 standard; cDNA; 2415 BP.  
XX  
AC AAF81819;  
XX  
DT 05 JUN 2001 (first entry)  
XX  
DE Human; G protein-coupled receptor protein A027 encoding cDNA.  
XX  
KW Human, G protein-coupled receptor, A027, nootropic, neuroprotective;  
KW hypotensive, antileukemic, antiarrhythmic, cardiac, antidiabetic,  
KW abortifacient, gene therapy; Alzheimer's disease, hypertension;  
KW pregnancy termination, rheumatism; allergy; angina pectoris; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 354..1649  
FT /\*tag= a  
FT /product= "A027"  
FT /note= "G protein coupled receptor"































Db 961 GGATTTCCAACTCCATCTGTAATCCCAATTCCTCTATGCAATTATCAATCAAAATTCATA 1020  
 QY 1374 AAAAATGTTTGTCTGAGTTTGTATTGTCATAGTAATAAATCAATTTCTCTCCAGACAA 1433  
 Db 1021 AAAAATGTTTGTCTGAGTTTGTATTGTCATAGTAATAAATCAATTTCTCTCCAGACAA 1080  
 QY 1434 AGGCATGCAATTCAGCAATTTACAATGATCGGCAAGAAAGATTTTCCCTCCAGAGAS 1493  
 Db 1081 AGGCATGCAATTCAGCAATTTACAATGATCGGCAAGAAAGATTTTCCCTCCAGAGAS 1140  
 QY 1494 AATCCAGTGGAGAAATCAAGAGAGAAATATTGATGATGATGATGATGATGATGATGATG 1553  
 Db 1141 AATCCAGTGGAGAAATCAAGAGAGAAATATTGATGATGATGATGATGATGATGATGATG 1200  
 QY 1554 TGTGAACAGACAG 1613  
 Db 1201 TGTGAACAGACAG 1260  
 QY 1614 CTGGCTCAGAAATTCCTCTTTAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1649  
 Db 1261 CTGGCTCAGAAATTCCTCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296

## RESULT 12

ABK86292

ID ABK86292 standard; cDNA; 1737 BP.

XX AC ABK86292;

XX AC ABK86292;

DT 27-AUG-2002 (first entry)

XX DE Mouse TGR346a cDNA.

XX KW Mouse; TGR346a; gene; ss: G-protein coupled receptor; GPCR; TGR;

XX KW TGR-associated disorder; signal transduction; renal failure; nephritis;

XX KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;

XX KW diabetes insipidus; hyperprolactinemia; thirst disturbance; appetite;

XX KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;

XX KW circadian rhythm.

XX OS Mus sp.

XX PH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 1..1302

XX FT /\*tag= a

XX FT /product= "Mouse TGR346a protein"

XX FT

XX PN WO200242458-A2.

XX PD 30-MAY-2002.

XX PF

XX PF 21-NOV-2001; 2001WO-US43404.

XX PR 22-NOV-2000; 2000US-252841P.

XX PR 22-DEC-2000; 2000US-257636P.

XX PR 12-JAN-2001; 2001US-261377P.

XX PR 28-MAR-2001; 2001US-279554P.

XX PR 29-MAR-2001; 2001US-280696P.

XX XX

XX PA (TULAR) TULARIK INC.

XX XX

XX PI Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;

XX XX

XX WPI; 2002-463633/49.

XX DR P-PSDB; AAU97740.

XX XX

XX PT New isolated G-protein coupled receptor polypeptide, termed TGR, for

XX PT diagnosis and treatment of diseases such as renal failure, nephritis,

XX PT hypothyroidism, diabetes insipidus, and disturbances of thirst and

XX PT sleep

XX PS

XX PS Claim 22; Page 77-78; 98pp; English.

XX XX

XX XX

CC The invention relates to a G-protein coupled receptor polypeptide (GPCR),  
 CC termed TGR, and its associated nucleic acid. The sequences of the  
 CC invention are useful for identifying a compound that modulates the  
 CC transduction and for identifying a mammal having a TGR-associated  
 CC disorder. The proteins and nucleic acids are useful in diagnosis and  
 CC treatment of diseases or conditions such as renal failure, nephritis,  
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,  
 CC diabetes insipidus, hyperprolactinemia and disturbances of thirst,  
 CC sleep, temperature regulation, appetite, blood pressure or any other  
 CC syndrome or disease associated with the hypothalamus. The sequences can  
 CC be used in regulation of circadian rhythms, for use as genetic markers  
 CC for the identification of mutations associated with diseases associated  
 CC from GPCR inactivation in particular cell types and for identification of  
 CC modulators of GPCR signal transduction. This sequence represents cDNA  
 CC encoding the mouse TGR346a polypeptide.  
 XX

SQ Sequence 1737 BP; 473 A, 384 C, 392 G, 488 T, 0 other;

## Query Match

Best Local Similarity 41.2%; Score 994.6; DB 24; Length 1737.

Matches 1172; Conservative 0; Mismatches 253; Indels 12; Gaps 1;

QY 354 ATGTAAGAGATTAAATTATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 413  
 Db 1 ATGCAGGCGCTCAACATCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
 QY 414 ACGGGAG 473  
 Db 61 ACTGGGAG 120  
 QY 474 CCGGAG 533  
 Db 121 ACGGGAG 180  
 QY 534 TTGGCAATGCTCTGGTGTGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 593  
 Db 181 TTGGCAATGCTCTGGT 240  
 QY 594 AACATCTTTATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 653  
 Db 241 AACATCTTTATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 300  
 QY 654 CTCACCATGCTCTGAG 713  
 Db 301 CTCACCATGCTCTGAG 360  
 QY 714 GTGCCATTGTGCTCAGTCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 773  
 Db 361 GTGCCATTGTGCTCAGTCTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
 QY 774 GTGGAAAGGAG 833  
 Db 421 GTGGAAAGGAG 480  
 QY 834 AGGCTTTTCAATGCTAGT 893  
 Db 481 AGGCTTTTCAATGCTAGT 540  
 QY 894 TGGCACTGTGCAAAATGAGATCAAAATATCAAAATATCAAAATATCAAAATATCAAAATAT 953  
 Db 541 TGGCACTGTGCAAAATGAGATCAAAATATCAAAATATCAAAATATCAAAATATCAAAATAT 600  
 QY 954 TGGTTAGAAG 1013  
 Db 601 TGGTTAGAAG 660  
 QY 1014 ATCCCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1073  
 Db 661 ATCCCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 720  
 QY 1074 CTTTGGATAAAG 1133  
 Db 721 CTTTGGATAAAG 780









QY 901 TCGAACAACTTCAGATCAATATGACATCCCTATATGAAAGGAACACATCTGCTGCTTAG 960  
 Db 61 TCGAACAACTTCAGATCAATATGACATCCCTATATGAAAGGAACACATCTGCTGCTTAG 120  
 QY 961 AAGAGTGGACGAGCTCTGACAGGAGAGATCTACATGCTTATCCTTGTCTATCTCT 1020  
 Db 121 AAGAGTGGACGAGCTCTGACAGGAGAGATCTACATGCTTATCCTTGTCTATCTCT 180  
 QY 1021 TCTCTCTGCTCTTATGCTGATGCTTATCTCTGATGCTTATCTCTGATGCTTATCTCT 1080  
 Db 181 TCTCTCTGCTCTTATGCTGATGCTTATCTCTGATGCTTATCTCTGATGCTTATCTCT 239  
 QY 1081 TAAAGAAAGAGTTGGGATGCTTCCAGTCTTCCAGTCTTCCAGTCTTCCAGTCTTCC 1140  
 Db 240 TAAAGAAAGAGTTGGGATGCTTCCAGTCTTCCAGTCTTCCAGTCTTCCAGTCTTCC 299  
 QY 1141 AATAGCTAGTAAAGAAAGAGTCTTATCTCTGATGCTTATCTCTGATGCTTATCTCT 1200  
 Db 300 AATAGCTAGTAAAGAAAGAGTCTTATCTCTGATGCTTATCTCTGATGCTTATCTCT 359  
 QY 1201 CTGTGCTGGGACCATTCCTGATGCTTGTCCATATGATGATGATGATGATGATGATG 1260  
 Db 360 CTGTGCTGGGACCATTCCTGATGCTTGTCCATATGATGATGATGATGATGATGATG 419  
 QY 1261 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 Db 420 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479  
 QY 1321 CCAACTCCATCTGTAATCCCATGCTCTATGCA 1352  
 Db 480 CCAACTCCATCTGTAATCCCATGCTCTATGCA 511

Search completed: June 17, 2003, 08:24:24  
 Job time : 1368 secs



Genome version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 07:57:49 ; Search time 3376 seconds  
(without alignments)  
11585338 Million cell updates/sec

Title: us-10-070-241B-2

Perfect score: 2415

Sequence: 1 gccagagccagcagccct... aaaaaaaaaaaaaaaaaaag 2415

Scoring table: IDENTITY\_NUC  
Gapop 10 0 , Gapext 1 0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32306132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estlin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpi.\*

7: em\_estro.\*

8: em\_hic.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hic.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: gb\_gss.\*

18: em\_gss\_hum.\*

19: em\_gss\_inv.\*

20: em\_gss\_pln.\*

21: em\_gss\_vrt.\*

22: em\_gss\_fun.\*

23: em\_gss\_mam.\*

24: em\_gss\_mus.\*

25: em\_gss\_other.\*

26: em\_gss\_pro.\*

27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 651.2 | 27.0        | 769    | 12 | BG471265 602512391 |
| 2          | 531.2 | 22.0        | 740    | 13 | B6729969 603350846 |
| 3          | 512.6 | 21.2        | 657    | 10 | B656182 HH656182   |
| 4          | 507.2 | 21.0        | 730    | 12 | HG169612 602321266 |
| 5          | 493   | 20.4        | 630    | 10 | BB649191 BB649191  |
| 6          | 478.2 | 19.8        | 636    | 10 | BB663258 BB663258  |

|    |       |      |     |    |          |
|----|-------|------|-----|----|----------|
| 7  | 432.4 | 17.9 | 645 | 10 | BB626475 |
| 8  | 380.8 | 15.8 | 651 | 10 | BB641290 |
| 9  | 374.2 | 15.5 | 634 | 10 | HH642180 |
| 10 | 370.8 | 15.4 | 866 | 14 | HQ735771 |
| 11 | 315.6 | 13.1 | 664 | 10 | HH044541 |
| 12 | 309   | 12.8 | 432 | 9  | AI307658 |
| 13 | 309   | 12.8 | 432 | 9  | AI308124 |
| 14 | 302.6 | 12.5 | 642 | 9  | AL651383 |
| 15 | 285   | 11.8 | 546 | 17 | A2454676 |
| 16 | 252.6 | 10.5 | 604 | 10 | AV760466 |
| 17 | 247.8 | 10.3 | 504 | 17 | AQ215853 |
| 18 | 246.2 | 10.2 | 360 | 17 | B91708   |
| 19 | 246.2 | 10.2 | 492 | 14 | PM993628 |
| 20 | 246.2 | 10.2 | 669 | 17 | AG155696 |
| 21 | 246   | 10.2 | 876 | 17 | AQ746417 |
| 22 | 245.8 | 10.2 | 672 | 17 | AG084721 |
| 23 | 244.2 | 10.1 | 921 | 17 | AQ738768 |
| 24 | 243.8 | 10.1 | 465 | 9  | AI613280 |
| 25 | 242.8 | 10.1 | 630 | 17 | AQ628459 |
| 26 | 242.6 | 10.0 | 524 | 17 | AQ278931 |
| 27 | 241.6 | 10.0 | 664 | 17 | AG040130 |
| 28 | 241.4 | 10.0 | 660 | 17 | AG113120 |
| 29 | 241.4 | 10.0 | 687 | 17 | AG113667 |
| 30 | 241.2 | 10.0 | 342 | 10 | AW276435 |
| 31 | 241.2 | 10.0 | 434 | 17 | AQ483420 |
| 32 | 241.2 | 10.0 | 924 | 17 | AQ748907 |
| 33 | 241   | 10.0 | 471 | 17 | AQ583362 |
| 34 | 240.8 | 10.0 | 381 | 12 | HC272215 |
| 35 | 240.8 | 10.0 | 444 | 17 | AQ347487 |
| 36 | 240.8 | 10.0 | 613 | 17 | AQ074754 |
| 37 | 240.6 | 10.0 | 444 | 17 | B89781   |
| 38 | 240.6 | 10.0 | 868 | 12 | BF337291 |
| 39 | 240.6 | 10.0 | 877 | 17 | AQ740209 |
| 40 | 240.4 | 10.0 | 477 | 17 | AQ133650 |
| 41 | 240.4 | 10.0 | 533 | 17 | AQ082371 |
| 42 | 240.4 | 10.0 | 679 | 17 | AG167348 |
| 43 | 240.4 | 10.0 | 721 | 14 | BM679639 |
| 44 | 240.2 | 9.9  | 646 | 10 | BB27215  |
| 45 | 240   | 9.9  | 509 | 17 | AQ482472 |

ALIGNMENTS

RESULT 1  
BG471265 602512391F1 NIH\_MGC\_16 Homo sapiens cDNA clone IMAGE:4534903 5',  
LOCUS BG471265 769 bp mRNA linear EST 21-MAR-2001  
DEFINITION mRNA Sequence.  
ACCESSION BG471265  
VERSION BG471265.1 GI:13403540  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi,  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 769)  
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cqapbs-remail.nih.gov](mailto:cqapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: L10M1391 row: f column: 24  
High quality sequence stop: 765.  
Location/Qualifiers  
1..769  
/organism="Homo sapiens"















/note- Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', CAGACGAGAAGGATCTCAAGCATCTTTTCTTTTTVN 3']. cDNA was prepared by using trihalose thermo activated reverse transcriptase and subsequently attached for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subcloning to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GACGAGAGATCTTCGATTAAATAATATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified phagescript KS(+) after bulk excision from Lambda FLC I. -Retina RNA was provided by Stefano Gustlinch, Department of Neurobiology, Harvard Medical School, 220





```

ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata, Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
COMMENT      Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnlnl.gov/bbrp/image/image.html
Insert Length: 1414 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop. 419.
FEATURES
Source
1..432
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site: 1, Not 1, Site 2, Eco RI,
Plasmid DNA from the normalized library NCI-CGAP_Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT      129 a 92 c 92 g 119 t
ORIGIN
Query Match      12.8%; Score 309; DB 9; Length 432.
Best local Similarity 79.3%; Pred. No. 1.2e-21;
Matches 432; Conservative 0; Mismatches 0; Indels 113; Gaps 1;
226 -----
1131 GAAATGTCACAAATACCCAGGAAGAACACAGCTGTCATTATCATGCTGACATGGTG 1190
226 -----GGAGGAAGAAACAGCTGTCATTATGATGCTGACATGGTG 186
1191 GCTCTCTTTGCTGTGTGGGACCATTCCTGATGTTGTCATATGATGATGATGATGAT 1250
185 GCTCTCTTTGCTGTGTGGGACCATTCCTGATGTTGTCATATGATGATGATGATGAT 126

FEATURES
Source
1..432
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site: 1, Not 1, Site 2, Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT      129 a 92 c 92 g 119 t
ORIGIN
Query Match      12.8%; Score 309; DB 9; Length 432.
Best local Similarity 79.3%; Pred. No. 1.2e-21;
Matches 432; Conservative 0; Mismatches 0; Indels 113; Gaps 1;
226 -----
1131 GAAATGTCACAAATACCCAGGAAGAACACAGCTGTCATTATCATGCTGACATGGTG 1190
226 -----GGAGGAAGAAACAGCTGTCATTATGATGCTGACATGGTG 186
1191 GCTCTCTTTGCTGTGTGGGACCATTCCTGATGTTGTCATATGATGATGATGATGAT 1250
185 GCTCTCTTTGCTGTGTGGGACCATTCCTGATGTTGTCATATGATGATGATGATGAT 126

```



# JOURNAL COMMENT

plasmid inserts  
Unpublished (2000)  
Contact: Robert H. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0256 row. A column: 18  
Seq primer: CACACAGGAACAGCATCACC  
Class: plasmid ends  
High quality sequence stop: 546.

## FEATURES

source

1..546  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0256A18"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (g14732114gb1AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

180 a 95 c 107 g 164 t

Query Match 11.8%; Score 285; DB 17; Length 546;  
Best Local Similarity 74.0%; Pred No. 2.3e-19;

Matches 393; Conservative 0; Mismatches 125; Indels 13; Gaps 2;

QY 1266 TATGATGATGTCACATCAAGATGATTTTGGTCATCGTGCAAAATTATTGGATTTTCCAAC 1325  
DB 1 TATGATGATGTCACATCAAGA-GGTTTTTGGTGTGGCAAAACAAATTTGGCTTTTCAAC 59  
QY 1326 TCATCTGTAATCCATCTCTATCTATTATGAAAGAAAGCTTCAAAAAAATGTTTTG 1385  
DB 60 TCCATCTGTAATCCCTTTGTGTATGCATTATGAATGAAAGAACTTCAAAAGAAATTTTTTG 119  
QY 1386 TCTGCAGTTTGTATTGGCATAGTAATAAAGCTTCTCCACAGCAAGGCATGCAAT 1445  
DB 120 TCTGGGTTTGTATTGGCATAGTAAGAGAAACCTTCTCCACAGCAAGGCATGCAAT 179  
QY 1446 TCAGGAATTTACATGATCGGGAAGCAAGCAAGTTTCCCTCAGACAGAAATCCAGTGGAG 1505  
DB 180 TCTGGGATTTCAATGATGCAAGAGAGCAAGAAATTTATCAGCATCACAGCTCCAGTGGCG 239  
QY 1506 GAAACCAAGAGAAATTTATTTAGTATGTAAGATTGAAATTTGTAAGAGAGACA 1565  
DB 240 GAAGCCAAAGAGAGACTTATTCAGCATGCCAAGCTTGATGTCAAATTTGTGAGCAGCCA 299  
QY 1566 GAGGAGAGAAAAAGCAACGACATCTTGTCTCTTTTGGTCTCAACTGGCTGAGAT 1625  
DB 300 GGAGCAAGAAAGGCAACTCAAGGACAGAGCTTGGCTTCTTTAGTCTCAACTTCTGAAAC 359

QY 1626 TCTCTTTTAGACAGTGGGCATTAAATTATAAATAATCTTTCATAATTAATGCGCTTCAGAT 1685  
DB 360 TCTACTTTTCCGAGTGGACATGAACCTGTAATGATATCCCTCATAGCTAATATCATTTGTAT 419  
QY 1686 TGTAACCCCAAGAGAAAAATTTATTTGAGCAAGAGGTCAAAATACACTCTTTTAAATTCCTTAAGAT 1745  
DB 420 -----GCAAGAGTTATTTTAAAGCAAGGTCAAGACTATTTTTTTAAATGACCA 467  
QY 1746 GATCACAAGAGAAAAACAATCATGTTCATTAATAAATGACACAGGCT 1796  
DB 468 GAAAGAAANAAACATGTTTTTCAATTAAATGAATAAATACATACACT 518

Search completed: June 17, 2003, 11:04:45  
Job time : 3392 secs







## RESULT 2

US-08-694-915-5  
 : Sequence 5, Application US/08694915  
 : Patent No. 5811545  
 : GENERAL INFORMATION:  
 : APPLICANT: Adamou, Julie  
 : APPLICANT: Kirkpatrick, Robert  
 : APPLICANT: Rosenberg, Martin  
 : TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE  
 : NUMBER OF SEQUENCES: 7  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: SmithKline Beecham Corporation  
 : STREET: 709 Swedeland Road  
 : CITY: King of Prussia  
 : STATE: PA  
 : COUNTRY: USA  
 : ZIP: 19406-2799  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette  
 : COMPUTER: IBM Compatible  
 : OPERATING SYSTEM: DOS  
 : SOFTWARE: FastSeq Version 1.5  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/694,915  
 : FILING DATE:  
 : CLASSIFICATION:  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER:  
 : FILING DATE:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Han, William T.  
 : REGISTRATION NUMBER: 34,344  
 : REFERENCE/DOCKET NUMBER: ATG50017  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 610-270-5219  
 : TELEFAX: 610-270-5090  
 : TELEX:  
 : INFORMATION FOR SEQ ID NO: 5:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 3742 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : HYPOTHETICAL: NO  
 : ANTI-SENSE: NO  
 : FRAGMENT TYPE:  
 : ORIGINAL SOURCE:  
 : US-08-694-915-5

Query Match 9.9%; Score 238.4; DB 1; Length 3742;  
 Best Local Similarity 90.8%; Pred. No. 5.8e-41;  
 Matches 265; Conservative 0; Mismatches 26; Indels 1; Gaps 1;  
 QY 1909 AGCCAGGACAGGAGGCTCATGCTGTAATGCCAGCAATTTGGAGAGGTTAGAGGGGGG 1968  
 DB 2321 AGCGGGGAGGAGGTTGATCAAGTGTAAATTTGAGAGGTTGAGAGGAGGAGGAGG 2380  
 QY 1969 ATCAGGAGTACAGAGATCAAAATCATCTGATTAATAGGTTAAAGGATATCTCTGCTA 2028  
 DB 2481 ATCAGGAGTACAGAGATCAAAATCATCTGATTAATAGGTTAAAGGATATCTCTGCTA 2440  
 QY 2029 AAAATATG AAAATATGAGGAGGAGTGTGAGGAGAAATGTGATGAGGAGTACTTTGAG 2087  
 DB 2441 AAAATATG AAAATATGAGGAGGAGTGTGAGGAGAAATGTGATGAGGAGTACTTTGAG 2500  
 QY 2088 GCTCAGGAGGAGATATCTGCTGATTAATAGGTTAAAGGATATCTCTGCTGATC 2147  
 DB 2501 GCTCAGGAGGAGATATCTGCTGATTAATAGGTTAAAGGATATCTCTGCTGATC 2560  
 QY 2148 GCACTGACTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2199  
 DB 2561 GCACTGACTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2612

## RESULT 3

US-08-724-394A-20  
 : Sequence 20, Application US/08724394A  
 : Patent No. 5872237  
 : GENERAL INFORMATION:  
 : APPLICANT: Feder, John N.  
 : APPLICANT: Kronmal, Gregory S.  
 : APPLICANT: Lauer, Peter M.  
 : APPLICANT: Ruddy, David A.  
 : APPLICANT: Thomas, Winston  
 : APPLICANT: Tsuchihashi, Zenta  
 : APPLICANT: Wolff, Roger K.  
 : TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
 : TITLE OF INVENTION: Sequences and Antibodies Thereof  
 : NUMBER OF SEQUENCES: 31  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
 : STREET: Two Embarcadero Center, 8th Floor  
 : CITY: San Francisco  
 : STATE: CA  
 : COUNTRY: USA  
 : ZIP: 94111-3834  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/724,394A  
 : FILING DATE: 01-OCT-1996  
 : CLASSIFICATION: 536  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Fitts, Renee A.  
 : REGISTRATION NUMBER: 35,136  
 : REFERENCE/DOCKET NUMBER: 017957-000100  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 415-576-0200  
 : TELEFAX: 415-576-0300  
 : INFORMATION FOR SEQ ID NO: 20:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 246240 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: not relevant  
 : TOPOLOGY: not relevant  
 : MOLECULE TYPE: cDNA  
 : FEATURE:  
 : NAME/KEY: misc.feature  
 : LOCATION: 1..246240  
 : OTHER INFORMATION: /note= "HILA-II-CONTIG"  
 : US-08-724-394A-20

Query Match 9.9%; Score 238.2; DB 2; Length 246240;  
 Best Local Similarity 86.2%; Pred. No. 2.8e-40;  
 Matches 275; Conservative 0; Mismatches 43; Indels 1; Gaps 1;  
 QY 1895 AAAGCAACATGGAAGGCGGAGGACGAGTGTGATGATGATGATGATGATGATGATGATG 1954  
 DB 2974 AAAAGCTCTTTATAGGG 3033  
 QY 1955 GCGTACGCGGGGATCAGGAGTCAAGAGATCAAAACCATGCTGGCTAGACACGCTGAAA 2014  
 DB 3034 GAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3093  
 QY 2015 CCGCATCTCTGCTAAATATAC-AAAAATTAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGG 2073  
 DB 3094 CCGCGTCTCTACTAAAAATACAAAAATTTAGCGGGGGGGGGGGGGGGGGGGGGGGGGGG 3153  
 QY 2074 CATATATTGAGAGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2133  
 DB 3154 CAGCTACTTGGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3213







[illegible]

```

RESULT 11
US-09-128-155-16/c
? Sequence 16, Application US/09128155
? Patent No. 6117654
? GENERAL INFORMATION:
? APPLICANT: Pan, Yaog
? TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
? TITLE OF INVENTION: AND USES THEREOF
? FILE REFERENCE: 09404/052001
? CURRENT APPLICATION NUMBER: US/09/128,155
? CURRENT FILING DATE: 1998-08-04
? EARLIER APPLICATION NUMBER: US 60/091,650
? EARLIER FILING DATE: 1998-07-02
? EARLIER APPLICATION NUMBER: US 60/054,646
? EARLIER FILING DATE: 1997-08-04
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 16
? LENGTH: 152331
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(152331)
? OTHER INFORMATION: n - A,T,C or G
US-09-128-155-16

```

[illegible]

RESULT 12  
US-09-128-155-17  
; Sequence 17, Application US/09128155

```

: Patent No.: 6117654
: GENERAL INFORMATION:
: APPLICANT: Pan, Yang
: TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: 09404/052001
: CURRENT APPLICATION NUMBER: US/09/128,155
: CURRENT FILING DATE: 1998-08-03
: EARLIER APPLICATION NUMBER: US 60/091,650
: EARLIER FILING DATE: 1998-07-02
: EARLIER APPLICATION NUMBER: US 60/054,646
: EARLIER FILING DATE: 1997-08-04
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 17
: LENGTH: 176373
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)..(176373)
: OTHER INFORMATION: n - A,T,C or G
US-09-128-155-17

Query Match          9.7%; Score 234.8; DB 3; Length 176373;
Best Local Similarity 85.1%; Pred. No. 1.3e-39;
Matches 274; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY      1879   ACTTCACATAGCCTTAAAAAAGCCAACATCGAAGGCCAGGCACGGTGCGCTCATGCGCTCTAAT 1938
        |||||
Db       128687  ACTATAGAGTAGCTTTNAGATCTAACAGATCTGTATTGGAGGAGGCGGCTCATGCTGTATAAT 128746

QY      1939   CCCACACTCTGGGAGGCGCTAGACGGGGCGGATACAGAGCTCAGGAGATCAAAACCATCTCT 1998
        |||||
Db       128747  CCCAGCACTCTGGGAGGCGGAGCGGGCGGATCACAAAGGTGAGGAGATCAAGACCATCTCT 128806

QY      1999   GGTAATACGCTGAAGATCTTATCTCTATANAATAA'AAAA-ATTAAAGAGAGAGCTGAGTGG 2057
        |||||
Db       128807  GCCTAACCGAGTGAACCCCATCTCTACTAAAATACAAATATTAGCCAGGCGTGGTGG 128866

QY      2058   CCGGCACCTTAGTCCCAGCTACTTGGAGCGCTCAGCGCGGCAATTGGTGTCAAACCCGGG 2117
        |||||
Db       128867  CGGAGATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 128926

QY      2118   AGSGGAGCTTGGCACTGATCTTCAGATCACTGATCTGATCTGATCTGATCTGATCTGATCTG 2177
        |||||
Db       129427  AGGTAAGCTTGGCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 129486

QY      2178   AGATCTCGGCTCTCAAAAAAAAA 2199
        |||||
Db       128987  AGACTCGGCTCTCAAAAAAAAAA 129008

```

```

RESULT 13
US-09-593-995-10
: Sequence ID, Application US/09593995
: Patent NO. 6406888
: GENERAL INFORMATION:
: APPLICANT: Conklin, Darrell C.
: APPLICANT: Gao, Zeren
: TITLE OF INVENTION: HELICAL CYTOKINE ZALPHA33
: FILE REFERENCE: 99-38
: CURRENT APPLICATION NUMBER: US/09/593,995
: CURRENT FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: 60/139,121
: PRIOR FILING DATE: 1999-06-14
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 10
: LENGTH: 20598
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-593-995-10

```



Oy 2193 AAAAAAA 2199  
|11111|  
bb 8921 AAAAAAA 8927

Search completed: June 17, 2003, 11:08:50  
Job time : 241 secs

























GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 17, 2003, 06:11:12, Search time 242 Seconds

(without alignments)  
2690.147 Million cell updates/sec

Title: US-10-070-241B-1

Perfect score: 2235

Sequence: 1 MQALNITPQFSRLLRDHNH... RHIALFRSLAFNSPLDS5H 431

Scoring table:

|                          |  |
|--------------------------|--|
| BLOSUM62                 |  |
| Xgapop 10.0, Xgapext 0.5 |  |
| Ygapop 10.0, Ygapext 0.5 |  |
| Fgapop 6.0, Fgapext 7.0  |  |
| Delop 6.0, Delext 7.0    |  |

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+p2n.model -DEV=xlh
-O/cgn2_1/USPTO.spool/US10070241/runat_10062003_095455_20883/app_query_fasta_1.583
-DB=published_applications_NA_QFMT-fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOKUP=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10070241 -CGN_1_1_146 -runat_10062003_095455_20883
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -USFBLOCK=100
-LOGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTNS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Query Match | Length | ID   | Description                            |
|------------|-------------|--------|------|--|
| 1          | 2235        | 100.0  | 1296 | 9 US-09-990-940-5 Sequence 5, Appli    |
| 2          | 2235        | 100.0  | 1296 | 10 US-09-826-508-21 Sequence 21, Appli |
| 3          | 2235        | 100.0  | 1296 | 10 US-09-899-532-3 Sequence 1, Appli   |
| 4          | 2235        | 100.0  | 1710 | 10 US-09-899-532-3 Sequence 3, Appli   |

#### ALIGNMENTS

RESULT 1

```
US-09-990-940-5
; Sequence 5, Application US/09990940
; Publication No. US0030027252A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiaganq
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: An, Songzhu
; APPLICANT: Dai, Kang
; APPLICANT: Gupta, Jamila S.
; APPLICANT: Talarik Inc.
; TITLE OF INVENTION: No. US0030027252A1el Receptors
; FILE REFERENCE: 018781-0074100S
; CURRENT APPLICATION NUMBER: US/09/990,940
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/252,841
; PRIOR FILING DATE: 2000 11 22
; PRIOR APPLICATION NUMBER: US 60/257,636
; PRIOR FILING DATE: 2000 12 22
; PRIOR APPLICATION NUMBER: US 60/261,377
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/279,554
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/280,696
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 54
```

Sequence 15, Appli  
Sequence 17, Appli  
Sequence 73, Appli  
Sequence 5, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 7, Appli  
Sequence 3, Appli  
Sequence 4, Appli  
Sequence 17, Appli  
Sequence 1, Appli  
Sequence 25, Appli  
Sequence 11438, A  
Sequence 1, Appli  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 38097, A  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 3, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 55, Appli  
Sequence 50, Appli  
Sequence 288, App  
Sequence 35, Appl  
Sequence 240, App  
Sequence 7, Appli  
Sequence 119, App  
Sequence 1, Appli



```
QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60
Db 121 CCGGACCGCCAGCTGGCCCTCGTCTCACCAGCGTGTCTATCTTCGCCCTGGCCCTC 180
QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetAlaArgThrValThr 80
Db 181 TTTGGCAATGCTCTGTTCTTATATGAGTACCTGATGATGATGATGATGATGATGATGAT 240
QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100
Db 241 AACATCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysLysMet 120
Db 301 CTCACCATGCTCCAGAACATTTCCACAACTGGCTGGGGGTGCTTTTCATTTGCAAGATG 360
QY 121 ValProPheValGlnSerThrAlaValValThrGluLeuLeuThrMetThrCysIleAla 140
Db 361 GTGGCAATTTGTCAGTCTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnIleThrAsnArg 160
Db 421 GTGGAAGGACACAGGACATGTCATCTCTTTTAAATGAAGTGGCAATACACCAACCGA 480
QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180
Db 481 AGGCTTTTCAAAAGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 181 TrpHisValGlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCys 200
Db 541 TGGACACGTGCAACAACTGAGATCAAAATATGATTTCTATATGAAAAAGAACATCTGC 500
QY 201 CysLeuGluGluTrpThrSerProValHisGlnLysIleCysTrpThrPheIleLeuVal 220
Db 601 TGTTCAGAGAGTGGACAGCCCTGTGCACAGAGATCTACACCACTTCATCTCTGTGTC 660
QY 221 IleLeuPheLeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGlu 240
Db 661 ATCTCTTCTCCCTCCCTCTTATGCTGATGCTTATCTGACAGTAAATTTGGTTATGAA 720
QY 241 LeuTrpIleLysLysArgValGlyAspGlySerValLeuArgThrIleHisGlyLysGlu 260
Db 721 CTTTGGATAAAGAAAAGATTGGGGATGGTTCAGTGTCTGAACTATTCATGGAAGAA 780
QY 261 MetSerLysIleAlaArgLysLysArgAlaValIleMetMetValThrValAla 280
Db 781 ATGTCCTCAAAATAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 281 LeuPheAlaValCysTrpAlaProPheHisValValHisMetMetIleGluTyrSerAsn 300
Db 841 CTCTTTGCTGTGCTGGCCACCATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 301 PheGluLysGluTyrAspAspValThrIleLysMetIlePheAlaIleValGlnIleIle 320
Db 901 TTTCAAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 321 GlyPheSerAsnSerIleCysAsnProIleValTyrAlaPheMetAsnGluAsnPhelLys 340
Db 961 GGATTTTCCAACTCCATCTGTAATCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 341 LysAsnValLeuSerAlaValCysTyrCysIleValAsnLysThrPheSerProAlaGln 360
Db 1021 AAAAATGTTTGTCTGCAATTTGTTATTTGATAGTAAATAAAACCTTCTCTCCAGCAAA 1080
QY 361 ArgHisGlyAsnSerGlyIleThrMetMetArgLysLysAlaLysPheSerLeuArgGlu 380
Db 1081 AGGCATGGAATTCAGGAATTTACATGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 381 AsnProValGluThrLysGlyAlaPheSerAspGlyAsnIleGluValLysLeu 400
Db 1141 AATCCAGTGGAGGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
```

---

```
QY 401 CysGluGlnThrGluGluLysLysLysLeuLysArgHisLeuAlaLeuPheArgSerGlu 420
Db 1201 TGTGAACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 421 LeuAlaGluAsnSerProLeuAspSerGlyHis 431
Db 1261 CTGGCTGAGAAATCTCTCTTTAGACAGTGGGCAT 1293
RESULT 3
US-09-899-532-1
; Sequence 1, Application US/09899532
; Patent No. US20020048791A1
; GENERAL INFORMATION:
; APPLICANT: Bloomquist, Brian T.
; APPLICANT: Zhelmin, Leonid
; TITLE OF INVENTION: Human Neuropeptide Y-like G
; FILE OF INVENTION: Protein-Coupled Receptor
; FILE REFERENCE: 02973.00040
; CURRENT APPLICATION NUMBER: US/09/899,532
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/41b,523
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-899-532-1
Alignment Scores:
Pred. No.: 2,79e-259 Length: 1296
Score: 2235.00 Matches: 431
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-10-070-241b-1 (1-431) x US-09-899-532-1 (1-1296)
QY 1 MetClnAlaLeuAsnIleThrProGluGlnPheSerArgLeuLeuArgAspHisAsnLeu 20
Db 1 ATCAGAGCGCTTAACATTAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 21 ThrArgGluGlnPheIleAlaLeuTyrArgLeuArgProLeuValTyrThrProGluLeu 40
Db 61 ACAGCGGAGAGAGTTCATCTGCTGTACCGGCTGGAGCGCTGCTGTACACCCAGAGCTG 120
QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60
Db 121 CCGGACCGCCAGCTGGCCCTCGTCTCACCAGCGTGTCTATCTTCGCCCTGGCCCTC 180
QY 61 PheGlyAsnAlaLeuValPheTyrValValThrArgSerLysAlaMetArgThrValThr 80
Db 181 TTTGGCAATGCTCTGTTCTTATATGAGTACCTGATGATGATGATGATGATGATGATGATG 240
QY 81 AsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIlePro 100
Db 241 AACATCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 101 ValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGlyAlaPheIleCysLysMet 120
Db 301 CTCACCATGCTCCAGAACATTTCCACAACTGGCTGGGGGTGCTTTTCATTTGCAAGATG 360
QY 121 ValProPheValGlnSerThrAlaValValThrGluLeuLeuThrMetThrCysIleAla 140
Db 361 GTGGCAATTTGTCAGTCTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 141 ValGluArgHisGlnGlyLeuValHisProPheLysMetLysTrpGlnIleThrAsnArg 160
Db 421 GTGGAAGGACACAGGACATGTCATCTCTTTTAAATGAAGTGGCAATACACCAACCGA 480
QY 161 ArgAlaPheThrMetLeuGlyValValTrpLeuValAlaValIleValGlySerProMet 180
```









|    |     |  |     |
|----|-----|--|-----|
| Db | 126 | "AAATTTTGAAGGAATATGATGATGTCACAAATCAAGATGATTTTTCGTATCTGCAAA"      | 67  |
| QY | 319 | eIleGlyPheSerAsnSerIleCysAsnProIleValIyrAlaPheMetAsnGluAsnPh     | 339 |
| Lb | 66  | TATTGSAATTTTCAACTGCATCTGATCTAAATCCCAATGTCTATGCAATTATGAATGAAAACIT | 7   |
| QY | 339 | eIys 340   |     |
| Db | 6   | CAAA 3   |     |

```

RESULT 8
US-09-866-248A-5
/ Sequence 5, Application US/09866248A
/ Publication No. US20020198367A1
/ GENERAL INFORMATION:
/ APPLICANT: Gerald, Christophe P.G.
/ APPLICANT: Jones, Kenneth A.
/ APPLICANT: Bonini, James A.
/ APPLICANT: Borowsky, Beth
/ TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide RF (NPFF) Receptors
/ TITLE OF INVENTION: and Uses Thereof
/ FILE REFERENCE: 1795/57155-A
/ CURRENT APPLICATION NUMBER: US/09/866,248A
/ CURRENT FILING DATE: 2001.05.25
/ PRIOR APPLICATION NUMBER: 09/161,113
/ PRIOR FILING DATE: 1998-09-25
/ NUMBER OF SEQ ID NOS: 42
/ SOFTWARE: PatentIn Ver. 2.0 - beta
/ SEQ ID NO 5
/ LENGTH: 1302
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-866-248A-5

```

```

US-09-866-248A-5

Alignment Scores:
Pred. No.:
3e-50      Length:      1302
Score:      502.00     Matches:      128
Percent Similarity: 50.54%  Conservatives: 87
Best Local Similarity: 30.12% Mismatches: 162
Query Match: 22.46%  Indels:      48
DB:          9        Gaps:        12

US-10-070-241B-1 (1-431) x US-09-866-248A-5 (1-1302)

```

US-10-070-241H-1 (1-431) x US-09-866-248A-5 (1-1302)

|    |     |  |     |
|----|-----|--|-----|
| QY | 14  | LeuLeuArgAspHisAsnLeuThrArgGluGlnPheIleAlaLeuTyrArgLeuArgPro | 33  |
|    |     |  |     |
| Db | 106 | CTGTACTCAGATAATTAATTACC-----TATGTGAAC TACTATCTTCACAGCCT      | 156 |
| QY | 34  | LeuValTyrThrProGluLeuProGlyArgAlaIysLeuAlaLeuValLeuThrGlyVal | 53  |
|    |     |  |     |
| Db | 157 | CAAGTGT-----GCACCAATCTTCATTATTTCCTACTTT                      | 189 |
| QY | 54  | LeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeuValPheTyrValValThrArgSer | 73  |
|    |     |  |     |
| Db | 190 | CTGATCTCTCTTTTGTGTCATGATGGAAATACTGTCTTTCCTTTATTGTAAATCAGGAAC | 249 |
| QY | 74  | LysAlaMetArgThrValThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeu | 93  |
|    |     |  |     |
| Db | 250 | AAACATATGACACAGTACTACTATCTCTTATTATAAATTGGCATAGTGAATTTACTAT   | 309 |
| QY | 94  | IleThrPhePheCysIleProValThrMetLeuGlnAsnIleSerAspAsnTrpLeuGly | 113 |
|    |     |  |     |
| Db | 310 | CTTGGCATATTCTGCATGCTCTATAACATGCTGTGCACAATATTATAGCAGATGGCCATT | 369 |
| QY | 114 | GlyAlaPheIleCysLysMetValProPheValGlnSerThrAlaValValThrGluIle | 133 |
|    |     |  |     |
| Db | 370 | GGAAACACCATCTGCACATCAGTGGATTGTCTGCAGGCAATATCTCTCCAGCTTTCAGTC | 429 |
| QY | 134 | LeuThrMetThrCysIleAlaValGluArgHisGlnGlyLeuValHisProPheLysMet | 153 |
|    |     |  |     |
| Db | 430 | TTTATGTTAGTGTGAATTTGATATATATGTTTATATGTTGTTTATATATTTTAAACCA   | 489 |
| QY | 154 | LysTrpGlnTyrThrAsnArgAlaPheThrMetLeuGlyValValThrPleuValAla   | 173 |







```

Db 529 AGCGTACGGTCACCGGAGGAGAACACACACATTCATGTGGACGCGGCAACCGGCGCCAC 588
QY 200 -----CysCysLeuGluGluThrThrSerProValHisGlyLeuLeuThrThr 216
Db 589 CCGCTCTACTCTCTGGGAGCCCTGGCGCCAAAGAGGCGATGCGCAGGGTCTACACAC 648
QY 217 PheLeuLeuValLeuPheLeuLeuProLeuMetValMetLeuLeuLeuLeuSer 236
Db 649 GTGCTCTCTCTGACATCTACTGCGCGGTGGCGCTCATCGTCTGATGATGATGATG 708
QY 237 IleGlyTyrGluLeuThrPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 256
Db 709 ATCGCGCGCAAGCTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 750
QY 257 HisGlyLeuGluMetSer-----LysIleAlaArgLysLysLysLysLysLys 274
Db 751 ---GCCGAGGAGCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 807
QY 275 MetValThrValValAlaLeuPheAlaValCysTyrAlaProPheHisValVal 294
Db 808 CTGGTCATGTTGGAGTCTTTCTTCAAGTGTCTGTGGTGGTGGTGGTGGTGGTGG 867
QY 295 MetIleGluTyrSerAsnPheGluLysGluTyrAspAspValThrIleLysMet 314
Db 868 CTGATCGACTACGGCGCAGCTCAGCGCGCGCGAGCTGCGACCTGGTGCACGCT 927
QY 315 AlaIleValGlnIleIleGlyPheSerAsnSerIleCysAsnProIleValVal 334
Db 928 CCGTTCGGCGGACATGGCTGGCTGCTTCAACAGCGCGCGCGCGCGCGCGCGCG 987
QY 335 MetAsnGluAsnPheLysLysAsnValLeuSerAla 346
Db 988 TTCAACGAGAACTTCGCGCGCGCTTCCAGCGCGCGC 1023

RESULT 13
US-10-090-569-4
; Sequence 4, Application US/10090569
; Publication No. US20030008329A1
; GENERAL INFORMATION:
; APPLICANT HINDMA, Shuji
; FUJII, Ryo
; FUKUSUMI, Shoji
; OHTAKI, Tetsuya
; HOSOKA, Masaki
; OHGI, Kazuhiro
; ONDA, Haruo
; TITLE OF INVENTION: CALANIN RECEPTOR PROTEIN, PRODUCTION AND USE THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,650
; FILING DATE: 11-OCT-1995
; APPLICATION NUMBER: 7-134412
; FILING DATE: 31-MAY-1995
; APPLICATION NUMBER: 6-326610
; FILING DATE: 28-DEC-1994
; APPLICATION NUMBER: 6-247599

```

```

; FILING DATE: 13-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S
; REGISTRATION NUMBER: 34,235
; REFERENCE/PROJECT NUMBER: 45901
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1044
; TYPE: Nucleic acid
; STRAND: Double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; IDENTIFICATION METHOD: S
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-090-569-4

Alignment Scores:
Pred. No.: 8,68e-42 Length: 1044
Score: 430.50 Matches: 101
Percent Similarity: 52.81% Conservative: 59
Best Local Similarity: 33.33% Mismatches: 110
Query Match: 19.26% Indels: 33
DB: 9 Gaps: 9

US-10-090-569-4 (1-431) x US-10-090-569-4 (1-1044)
QY 46 LeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeuPheGlyAsnAlaLeu 65
Db 103 ATTAGATGATAGTGTGTTGATCTGATTTTATGATGATGATGATGATGATGATG 159
QY 66 ValPheTyrValValThrArgSerLysAla-----MetArgThrValThrAsnIlePhe 83
Db 160 GTGATATACCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 219
QY 84 IleCysSerLeuAlaLeuSerAspLeuLeuIleThrPhePheCysIleProValThrMet 103
Db 220 ATACTATATCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 279
QY 104 LeuGlnAsnIleSerAspAsnTrpLeuGlyGlyAlaPheIleCysLysMetValProPhe 123
Db 280 ACCGTGTATGACATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 339
QY 124 ValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIleAlaValGluArg 143
Db 340 TCTTTCACCGTGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 399
QY 144 HisGlnGlyLeuValHisProPheLysMetLysTrpGlnTyrThrAsnArgAlaPhe 163
Db 400 TACTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
QY 164 ThrMetLeuGlyValValThrPheValAlaValIleValGlySerProMetThrPheVal 183
Db 460 CTGAGTGTGGTGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 184 GlnGlnLeuGluIleLysTyrAspPheLeuTyrGluLysGluHisIleCysCysLeuGlu 203
Db 520 CACGCTCTT-----TTCCATCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 564
QY 204 GluTrpThrSerProValHisGlnLysIleTyrThrThrPheIleLeuValIlePhe 223
Db 565 CAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624
QY 224 LeuLeuProLeuMetValMetLeuIleLeuTyrSerLysIleGlyTyrGluLeuTrpIle 243
Db 625 CTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY 244 LysLysArgValGlyAspGlySerValLeuArgThrIleIleGly-----LysGluMet 261
Db 664 -----GTCTTATATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 699

```



QY 319 IleIleGlyPheSerAsnSerIleCysAsnProIleValValtyrAlaPheMetAsnGluAsn 338  
 Db 1345 TGCCATGACCACTCCTCAGTGAACCCCATCATATATGCCCTTCTCTCAGAAAC 1404  
 QY 339 PheLysLys 341  
 Db 1405 TTCGGGAG 1413  
 RESULT 15  
 US-09-900-497-1  
 ; Sequence 1, Application US/09900497  
 ; Patent No. US2002008232A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Keith D.  
 ; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING NPV6-R  
 ; FILE REFERENCE: R-639  
 ; CURRENT APPLICATION NUMBER: US/09/900.497  
 ; CURRENT FILING DATE: 2001-07-06  
 ; PRIOR APPLICATION NUMBER: US 60/216,260  
 ; PRIOR FILING DATE: 2000-07-06  
 ; PRIOR APPLICATION NUMBER: US 60/221,474  
 ; PRIOR FILING DATE: 2000-07-27  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2281  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 US-09-900-497-1

Alignment Scores:  
 Pred No. 4 930-41 Length: 2281  
 Score: 429.00 Matches: 94  
 Percent Similarity: 52.65% Conservative: 95  
 Best Local Similarity: 26.18% Mismatches: 124  
 Query Match: 19.19% Indels: 46  
 DB: 10 Gaps: 11

US-10-070-241b-1 (1-431) x US-09-900-497-1 (1-2281)

QY 41 ProGlyArgAlaLysLeuAlaLeuValLeuThrGlyValLeuIlePheAlaLeuAlaLeu 60  
 Db 913 CCTTTTCTAGCCATCTCTTGCTACTCATAGTATATCTGATCTGATCTATCATCGGCATT 972  
 QY 61 PheGlyAsn---AlaLeuValPheTyrrValValThrArgSerIlyAlaMetArgThrVal 79  
 Db 973 TTGGAAACCTCTCTTATCATCATCATCTTTTAAAGAAACAGAGAGAGAGCTCAAAATGTT 1032  
 QY 80 ThrAsnIlePheIleCysSerLeuAlaLeuSerAspLeuLeuIleThrPheCysIle 99  
 Db 1033 ACCAAGATATCTGATTCACACCTGCTCCCTCTCTGACATCTTGCTGTCTCATGTGCATC 1092  
 QY 100 ProValThrMetLeuGlnAsnIleSerAspThrLeuGlyGlyAlaPheIleCysLys 119  
 Db 1093 CCTTTTACGGCTCATTAATCTGATGGACACTGGGTATTTGGGAACACTATGTGTAA 1152  
 QY 120 MetValProPheValGlnSerThrAlaValValThrGluIleLeuThrMetThrCysIle 139  
 Db 1153 CTCACCTCTAGCTGAAATGCTCTGATTTCTGTCCTCATATTCCTCTGTTGTTGATT 1212  
 QY 140 AlaValGluArgHisGlnGlyLeuValHisProPheLysMetLysThrPheGlnThrAsn 159  
 Db 1213 GCTATTGACGATATCAGCTGATTTGTAACCCG-----CGTGGCTGGAAACCCACAGTA 1266  
 QY 160 ArgArgAlaPheThrMetLeuGlyValValThrLeuValAlaValIleValGlySerPro 179  
 Db 1267 GCTCATGCTATTGAGGATCATCTTATTTGCTATTCTCTGACATGCTATATCC 1326  
 QY 180 MetThrHisValGlnGlnLeuGluIleLysTyrrAspPheLeuTyrrGluLysGluHis--- 198  
 Db 1327 TTATTC-----CTGTCTACCACTCAGCAATGAGGCTCTTCATAAT 1368

QY 199 -----IleCysCysLeuGluGluThrSer 207  
 Db 1369 CTCCTCTCCCTACATCATCTACACCCACAGATAGCTGCTGGGAGAAATTTGGCCCTCT 1428  
 QY 208 ProValHisGlnLysIleTyrrThrPheIleLeuValIleLeuPheLeuLeuProLeu 227  
 Db 1429 AAAGTGAACCAACTCTCTTTTCTACATCATTTATTTATGCTCCAGTATTTTGTCCCTCTG 1488  
 QY 228 MetValMetLeuIleLeuTyrrSerLysIleClyTyrrGluLeuThrPheLysLysArgVal 247  
 Db 1489 GCTTTCATCTTATCTGCTACCTCAAGATC-----GTTCTCTGCTCCGCAAAAGAA--- 1539  
 QY 248 GlyAspGlySerValLeuArgThrIleHisGlyLysGluMetSerLysIleAlaArgLys 267  
 Db 1540 -----ACTAGCCAGGAG 1587  
 QY 268 LysLysArgAlaValIleMetMetValThrValAlaLeuPheAlaValCysThrAla 287  
 Db 1588 AACAAAGAGAGTAAAGTATGTTGATTTTATGTTATGTTATGTTATGTTATGTTATGTT 1647  
 QY 288 PropheHisValValHisMetMetIleClyTyrrSerAsnPheGluLysGluTyrrAsp 307  
 Db 1648 CCTTGAACATTTTCAATGCTCACTTCACAGG-----TATCAACAG 1689  
 QY 308 ValThrIleLys-----MetIlePheAlaIleValGlnIleIleIleGlyPheSer 323  
 Db 1690 ATGCTGATGAGCTGCGACAG 1749  
 QY 324 AsnSerIleCysAsnProIleValTyrrAlaPheMetAsnGluAsnPheLysLysAsnVal 343  
 Db 1750 TCTACITGATATAATCTCTCTTTTATGATTTTCTCAACAAACAAACAAACAAACAAACAA 1809  
 QY 344 LeuSerAlaVal---CysTyrrCysIleValAsnLysThrPheSerProAlaGlnArg 361  
 Db 1810 ALGAGCTGATTCACACACTGTGTGGT-----GCTGACCTCAGAGAGAGAGAGAG 1854  
 QY 362 HisGlyAsnSerGlyIle---ThrMetMetArgLysLysAlaLysPheSerLeuArg 379  
 Db 1855 TATGAAATATTTGGCATGCTACTATGCACACAGATGAATGCAAGGGATCATTTAA 1911

Search completed: June 17, 2003, 08:01:50  
 Job time: 266 secs

